The Dissertation Committee for Vasudha Vasudevan Certifies that this is the approved version of the following dissertation:

Media Coverage of Mutual Funds

Committee:

Laura Starks, Supervisor

Andres Almazan

Keith Brown

Li Gan

Ron Kaniel

Media Coverage of Mutual Funds

by

Vasudha Vasudevan, B.S., M.B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2006

Dedication

To my family.

Acknowledgements

I would like to thank all members of my committee for their valuable comments and advice. I am immensely grateful to my dissertation chair, Laura Starks, for her invaluable mentorship and guidance throughout my Ph.D. program, and especially during the dissertation stage. Her support and counsel over the years have been vital to my progress and it has been an honor to work with her. I would also like to thank my family – my parents for always believing in me and motivating me to aim high, my daughter Sujani for filling my life with wonder and joy, and especially my husband Prasad for his support and understanding.

Media Coverage of Mutual Funds

Publication No._____________

Vasudha Vasudevan, Ph.D.
The University of Texas at Austin, 2006

Supervisor: Laura Starks

The principal focus of this dissertation is to investigate the role of media coverage in the investment decisions of mutual fund investors and the consequent effects on flows into the funds. I examine investor attention and learning effects by examining the relation between media coverage of mutual funds and the net investor flows to the funds. Using a database of nearly 10,000 news articles, I find that the existence and stance of media coverage affects net investor flows into the fund in ways consistent with investor attention and learning. Further, the media coverage does not have a uniform effect on flows. News articles with positive (negative) tones are associated with significant increases (decreases) in flows. I find that fund size and past performance influence the impact of media coverage on mutual fund flows. I also find that, as a fund ages and investors receive additional news about the fund, there are smaller effects from the news. This is consistent with the hypothesis that investors learn about funds through media coverage and that this knowledge affects their investment behavior. These results suggest that media coverage can have significant economic effects on mutual funds through the effects on investors' attention and learning.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Mutual funds have become very popular with investors over the past 25 years. Nearly half of all U.S. households owned mutual funds in 2006, up from less than 6% in 1980.[1] The total assets held by mutual funds increased by almost 6000% from $135 billion in 1980 to $8.9 trillion in 2005.[2] This explosive growth of the mutual fund industry has drawn considerable attention from the media as well as academic researchers. Much of the research has addressed issues of performance measurement and attribution, managerial behavior and incentives, economies of scale, fees and costs, and various other topics. One area that has received little attention is the impact of media coverage on mutual fund flows. The few studies that have focused on this topic have limited empirical work.

Nearly ninety percent of mutual fund investors are individuals.[3] Investors have thousands of mutual funds to choose from, and many investors have no training in how to select a mutual fund. When it comes to investment decisions, given the vast amount of information available and the limited resources, investors have to be selective in information processing. One potential important source of information and knowledge for investors is media coverage of mutual funds.

---

[1] Source: Investment Company Institute, 2006
[2] Source: Investment Company Institute, 2006
[3] Source: Investment Company Institute, 2006

The importance of media coverage as a source of investor information is highlighted in a survey of mutual fund investors conducted by the Office of the Comptroller of the Currency, the Securities and Exchange Commission, and a market research firm. In examining the responses to this survey, Alexander, Jones and Nigro (1998) find that in choosing which mutual funds to purchase, 42% of the investors state that they rely on financial publications like newspapers and magazines for information about the funds. Further, the authors find that for investors who purchase their funds directly from the fund company (versus through other sources such as brokers, employers, banks or insurance companies), 68% state that they rely on financial publications.

More than nine in ten households owning mutual funds had internet access in 2006 and of these, 57% used the internet to obtain investment information.[4] The scenario was very different a decade ago. The sample period for the data used in this dissertation is 1994-2000. Internet did not play as significant a role during that period. In the mid-nineties, most mutual funds websites were limited in terms of information and did not have much transaction capability. In fact, it was not until 2000 that Investment Company Institute conducted a study measuring mutual fund shareholders' access to the internet. Less than 32% of mutual fund shareholders visited fund-related websites in 2000 and 49% of shareholders who used the internet read financial publications online.[5] Thus, even at the end of the sample period, the media played a critical role in providing information about mutual funds to shareholders.

---

[4] Source: Investment Company Institute, 2006

[5] Source: Investment Company Institute, *Fundamentals* Vol.9, No.3

A few studies have touched upon the importance of media coverage. Merton (1987) mentions that a newspaper or other mass media story about a firm that reaches a large number of investors who are not currently shareholders, could induce some of this number to incur the set-up costs to follow and invest in the firm. Barber and Odean (2004) posit that news is a primary mechanism for catching investors' attention and provide evidence that investors buy stocks that catch their attention. Tetlock (forthcoming JF) suggests that measures of media content serve as a proxy for investor sentiment or non-informational trading. He finds that high levels of media pessimism robustly predict downward pressure on market prices, and unusually high or low values of media pessimism forecast high market trading volume. Barberis, Shleifer and Vishny (1998) state that current good news has power in predicting positive returns in the future. They highlight the need for a superior way of estimating the *strength* of news announcements. Sirri and Tufano (1998) find circumstantial evidence that garnering a larger share of current media citations is related to faster current growth of mutual funds. They conclude that "If the media are important determinants of consumer decisions, they probably deserve much more attention in the finance literature than our simple count ..".

This dissertation aims to contribute to the finance literature by conducting an in-depth analysis of the impact of media coverage, including the type and tone classifications of almost 10,000 news stories. The principal focus is an investigation of the pivotal role played by media coverage in the investment decisions of mutual fund investors and the consequent effects on flows into the fund. I propose that business news circulating in public can have significant, economically realizable and relevant information content. I further propose that this information can influence financial decisions by capturing the attention of investors and facilitating in their learning about mutual funds. This hypothesis can be tested empirically as the decisions of investors to purchase or redeem mutual fund shares should translate into net flows of capital into the fund. Specifically, I test the hypothesis that the existence of recent media coverage of a mutual fund influences investor flows into the fund. The results are consistent with the preliminary results in the previous studies mentioned above. Not only do I find that the existence of media coverage is associated with a significant increase in fund flows, I also find that such coverage has a stronger impact on fund flows than the fund's most recent performance. These results provide further evidence that media coverage can have significant economic effects on mutual funds.

The effects of media range from the obvious to the latent. A number of event studies provide evidence for the speed with which the market reacts to news.[6] However, the impact of media is not restricted to fast and directly measurable responses by stock prices. Media coverage of assets can have a multitude of latent effects, among which are learning and attention effects. These effects of media coverage, along with the relevant hypotheses, are described in greater detail in the following sections.

Learning Effects of Media Coverage

A salient feature of media coverage is that it is diverse, providing different types of information. The coverage may provide information about the fund, its characteristics

[6] For example, Patell and Wolfson (1984) and Jennings and Starks (1986) demonstrate that price movements in connection with dividend and earnings announcements through the Dow Jones News Service set in prior to publication. The main spike in stock prices follows within five to fifteen minutes after the publication. More impressively, sixty to ninety minutes later, price adjustments are for the most part concluded.

(such as fees, service, and distribution), its return performance, or insights into the managers' abilities. Consequently, the media coverage of mutual funds can have multiple effects on investors' learning about funds, and ultimately on their flows into or out of the funds. Investor learning from media coverage could also impact the fund's existing investors by leading them to reevaluate their current investments in such funds, resulting in either increased investments (in the case of a positive news story) or reduced investments (in the case of a negative news story).

Media coverage provides investors with the opportunity to learn about mutual funds and educates them to weigh various factors when buying or selling funds. Investing in mutual funds requires a learning process. There have been various approaches to learning in the finance literature. Some of the approaches focus on the process itself. For instance, many rational learning papers use Bayesian updating to model the learning process, while most behavioral finance papers on learning use psychological biases to model the investor learning process. Other approaches to learning focus on the source of the learning rather than the process. For example, investors can learn from their past purchase experience, or media coverage of assets can be a key source of investor learning.

A popular approach in the study of rational learning theory is to model investor learning using Bayesian updating. When investors have imperfect information about expected returns or cash flows, they must learn about the unknown price process from whatever information is available, which is formally modeled using Bayesian analysis. The underlying concept here is that the prices already "contain" the news and as a result, prices always represent an adequate reflection of fundamental values. Typically, in a model of rational learning, the investors (and sometimes the managers) update their posteriors on the basis of the history of observed returns as Bayesians. This method focuses primarily on the actual process of learning developed in these Bayesian models.

Several recent models for judging mutual fund performance and for examining the dynamics of the mutual fund flow-performance relation rely on investor learning about the mutual fund managers' skills through time (Berk and Green, 2004; Lynch and Musto, 2003; Pastor and Stambaugh, 2002; Baks, Metrick, and Wachter, 2001). In these models, investors learn about managerial ability by observing the realized excess returns the manager produces. This learning is then the source of the relationship between performance and the flow of funds.

The behavioral finance papers adopt an alternate perspective on learning. Hirshleifer (2001) argues that the purely rational approach is being subsumed by a broader approach based upon the psychology of investors. Empirical securities markets research in the last few decades has presented a body of evidence with systematic patterns that are not easy to explain with rational asset pricing models. He argues that the evidence on heuristics and biases seems to suggest that Bayesian updating is not fully descriptive of human behavior. In view of this, many recent studies have explicitly modeled how decision making occurs in a way that reflects psychological biases.

Daniel, Hirshleifer and Subrahmanyam (1998) focus on psychological evidence as a basis for assumptions about investor behavior. They theorize that news watchers rationally use fundamental news but ignore prices. Barberis, Shleifer and Vishny (1998) propose a parsimonious model of investor sentiment based on how investors form beliefs. The authors focus on learning about the time-series process of a performance measure such as earnings. Further, the authors refer to under-reaction evidence indicating that current good news has power in predicting positive returns in the future. While financial theorists have traditionally examined how information is transmitted by prices, volume, or corporate actions, Hirshleifer (2001) notes that media contagion of ideas and behavior also seems important.

Some studies approach learning by focusing on the source of the learning rather than the actual process of the learning method. One potential source of learning is past purchase experience of the investors. Barber, Odean and Zheng (2005) propose that front-end load fees and commissions are more obvious and salient; and front-end load fees are particularly salient for investors who have previously paid them. The authors provide evidence of learning among mutual fund investors where they find that experienced fund purchasers pay, on average, about half the front-end load fees of first time purchasers. From this evidence they conclude that investors have learned to avoid front-end load funds by experience.

Media coverage has been considered as a potential source of investor learning in some studies, albeit not about mutual funds. McQueen and Thorley (1997) find evidence that supports the investor learning hypothesis using data from the gold market. They conduct an empirical analysis of the predictive power of leading gold stock returns before and after the first time a gold analyst for Merrill Lynch was quoted in the *Wall Street Journal* about the lead/lag relationship between the prices of gold and the stocks of gold-producing companies. Consistent with the investor learning hypothesis, they find that the predictive power of the gold stock returns have diminished since the first public discussion of the anomaly.

In this dissertation, I focus on the media as the source of learning. Investors not only learn some facts about mutual funds but also learn how much importance to attach to those facts from the emphasis placed on them by the news media. Over the years, investors have had the opportunity to learn about mutual funds and to change the ways in which they weigh various factors when buying funds. The longer a fund has been around, the higher the potential for investor learning about the fund. Fund age is hence a good proxy for potential investor learning about the fund. I test the theory that investors have better knowledge of older funds that have been around for a longer time. I hypothesize that media coverage has a greater impact for younger funds than older mutual funds, and find the results consistent with the learning theory.

A key influence of media is knowledge gain leading to investor learning. In general, the impact of media is usually classified as cognitive, affective, or behavioral. Cognitive effects are those that concern the acquisition of information - what people learn, how beliefs are structured in the mind, how needs for information are satisfied or not. These effects include concerns about what is learned as well as how much is learned. I explore this aspect of learning empirically by examining the information content of the media coverage in terms of the tone or the posture of the news article, i.e., positive or negative slant in the article with regard to the fund. Consistent with the cognitive effects described above, I find that news articles with higher information content, that is, displaying a posture with regard to the fund, have a stronger impact on mutual fund flows than news articles with neutral tone or no posture.

Attention Effects of Media Coverage

Attention has long been recognized as a critical variable in the processing of mass media. Even if news is publicly available, it is not incorporated into investment decisions until investors pay attention. Huberman and Regev (2001) provide a vivid example. The publication of an article in the *New York Times* about a new cancer-curing drug from EntreMed attracted great public attention and generated a daily return of more than 300% in its stock, even though the same story had already been published more than five months earlier in *Nature* Magazine and in other newspapers.

There is no doubt that the media, as generators of attention, possess the potential to contribute to the overreaction in investor behavior by focusing public attention. The range and reach of the media has greatly increased as economic and non-economic events are today broadcast across borders and in real time. Schuster (2003) contends that the competitive situation on the media market intensifies the exaggeration of the contents and an increase in their emotional appeal through prominent placement and an eye-catching presentation of selected issues. Klibanoff et al. (1998) find that the price of a closed-end country fund reacts more strongly to news about its fundamentals when the country whose stocks the fund holds appears on the front page of the newspaper.

According to Griffin and Tversky (1992), in revising their forecasts, people focus too much on the strength of the evidence, and too little on its weight, relative to a rational Bayesian. The authors use a framework in which people update their beliefs based on the *strength* and *weight* of new evidence.[7] The Griffin and Tversky theory predicts that,

[7] *Strength* refers to such aspects of the evidence as salience and extremity, whereas *weight* refers to statistical informativeness, such as sample size.

holding the weight of information constant, news with more strength would generate a bigger reaction from investors. Asymmetrical price movements in response to comparable news give grounds to believe that investors overreact in some cases and under react in others. Barberis, Shleifer and Vishny (1998) provide an explanation, well-known to psychologists, that people attach too much significance to information that is particularly striking. According to the behavioral theory, conspicuous news events would lead to irrational price fluctuations, due to their increased visibility and the subsequent overreactions of many investors. In other words, the more visibly an event appears in the media, the stronger price reactions should be, regardless of how significant the event is in itself.

I test this prediction from behavioral theory in the context of media coverage of mutual funds. When a fund is portrayed with a numerical ranking in a news article, it is more likely to capture the attention of investors than a fund that does not appear in such an article.[8] I test for differences between the flows into funds that were listed in a ranking article against funds with almost identical returns that were not listed in the ranking article. I find significant differences in flows between the two sets of funds, thus providing evidence that there is a significantly stronger attention effect for the fund that is highlighted in the ranking article than the fund with almost identical returns that is not mentioned in the ranking article.

Merton (1987) advances the investor recognition hypothesis, in which he posits that since an individual investor will hold only those securities that he knows, familiarity with an asset can affect its value. The hypothesis suggests that investors conserve the

[8] For example, an article ranking the top three funds of the year.

resources required to gather information on stocks by actively following only a few stocks. Thus, each investor knows only about a subset of the available securities. Increased awareness of mutual funds due to media coverage implies that investors will view these funds as possibilities for their investment opportunity set in line with the investor recognition hypothesis.

Hirshleifer (2001) describes a strong and robust *mere exposure effect* in which exposure to an unreinforced stimulus tends to make people like it more. He suggests that the basis for this heuristic may be that what is familiar, being understood better, is often viewed as less risky. The implication derived here for media coverage of mutual funds is that investors are more likely to be familiar with funds that they recognize as a result of exposure to the fund in the media.

A key cognitive process involved in our experiencing media is attention. As a prerequisite to any other involvement with the media, we must select some information to attend to and process. In addressing the selective aspects of attention, Kahneman (1973) suggests that we *selectively attend* to some stimuli, or aspects of stimulation, in preference to others. Attended events are more likely to be perceived consciously and in greater detail. Kahneman further adds that recognition is highest for a stimulus which has all the critical features, is presented at high intensity, and is attended. The amount of attention that is allocated to a perceived article or event affects subsequent processing of the information contained in it.

Kahneman emphasizes that attention is a scarce cognitive resource. Odean (1999) posits that many investors limit their search to stocks that have recently captured their attention. Recent research suggests that investors find it extremely costly to search for mutual funds.[9] The constraints of time, resources and mobility impose severe limits on the acquisition of information. Thus, rather than incur the costs of finding a fund, they simply invest in funds that come to their attention. Sirri and Tufano (1998) classify media coverage about mutual funds as a measure of investor search costs. They contend that funds with more media coverage will grow faster and may have a stronger flow-performance relationship.

Barber and Odean (2004) propose that attention is a major factor determining the stocks that individuals buy. They provide evidence that individual investors are net buyers of attention-grabbing stocks, e.g., stocks in the news. Firms that are in the news are more likely to catch investors' attention than those that are not. I analyze the attention effect of media coverage on mutual funds by testing the hypothesis that mutual funds with higher numbers of news stories receive greater net inflows from investors. I find that in addition to fund flows responding to the *existence* of media coverage, the *frequency* of news stories also exerts significant influence on mutual fund flows. Historically, attention has been equated with simple exposure. However, research has brought to light that attention to news media appears to be a significant difference that accounts for substantial variation in learning beyond the act of simple exposure. In this dissertation, I test for the various aspects of attention by analysis of different classifications of the news articles based on strength, tone and type of the media coverage.

[9] Sirri and Tufano (1998), Huang, Wei and Yan (forthcoming JF)

Highlights of the results

In this dissertation, I examine several issues related to the media coverage of mutual funds. In particular, I examine the relation between a mutual fund's net flows and media coverage of that fund. I also examine the learning and attention effects via testing of hypotheses derived from relevant theories and literature. An advantage of examining attention and learning effects through the use of net fund flows is that one can obtain a more direct effect of news coverage than is possible with other types of assets, which have confounding valuation effects.

The database for media coverage of mutual funds comprising of nearly 10,000 news articles was developed specially for this study. I examine the determinants of media coverage using several univariate analyses and a multivariate probit estimation model. The results show that fund characteristics affect the likelihood of media coverage. Larger funds, those with extreme (high or low) performance, and those with more volatile returns are more likely to have news stories in general. Larger funds also have more positive news articles than smaller funds. I then investigate whether the existence and tone of media coverage affects investor attention and learning as is manifested by flows into the fund.

By examining the tone of the news articles, I test for differential effects on fund flows depending on whether the article has a positive, neutral or negative tone. I find small positive effects for the articles that have a neutral tone, i.e., there is no positive or negative slant in the article with regard to the fund, which suggest media coverage can have an attention effect on investors. In contrast, I find much stronger effects (in absolute magnitude) when the tone is positive or negative, suggesting that the news article also provides opportunity for learning by investors.

Napoleon is on record for saying that "Three hostile newspapers are more to be feared than a thousand bayonets." While this appears to have held true for politicians across time, I find that this also holds true for financial assets. I find that there is indeed such a thing as bad publicity. Media coverage does not have a uniform effect as the existence of news articles with a positive tone in a month is associated with a 1.5% increase in net investor flows. There is a 1% decrease in flows associated with the existence of news articles with a negative tone in a month. Considering that the average monthly fund flow for my sample is 1.08%, the above results suggest that media coverage can have significant economic effects on mutual funds.

I find a differential in flows based on fund age in the effects of a news story, which is driven by the younger funds. I find that, as a fund ages and investors receive additional news about the fund, there are smaller effects from the news. Older funds do not receive as high an inflow as do younger funds, even controlling for the differences in sizes between the funds. I also find that fund size and past performance influence the impact of media coverage on mutual fund flows. Media coverage seems to play a more significant role for smaller and lesser known funds, and the fund’s return performance enhances the impact of media coverage. I construct a test to differentiate between the attention and learning effects of media coverage on mutual fund flows. I examine the flows into funds that were listed in a ranking article against funds with almost identical returns that were not listed in the ranking article. I find significant differences in flows between the two sets of funds, suggesting that just getting into a ranking article can provide significant attention effects for a fund.

I employ the Heckman model of self-selection to conduct the empirical analyses. The study of the influence of media coverage on fund flows may have a potential endogeneity problem. It is highly unlikely that all mutual funds have an equal probability of receiving media coverage. It is also unlikely that the media randomly selects funds to write about. Selection bias may occur when funds that receive media coverage have a greater propensity to receive higher inflows. Then the positive effects of media coverage on mutual fund flows may be overstated. I control for this potential endogeneity by using the Heckman selection model to account for the selection bias.

I conduct an additional robustness check on the results using alternate specification of the regression model. I employ the Sirri and Tufano (1998) piecewise linear specification for return performance. I then use the Fama-MacBeth (1973) methodology by running monthly cross-sectional regressions and then averaging the coefficients across the 84 months using Newey-West (1987) t-statistics to test the significance of the coefficients. The results for these specifications are consistent with the original specification of the regression model. The economic and statistic significance of the news variables are thus robust for different model specifications. I also conduct Granger causality tests and find evidence in support of recent media coverage leading to mutual fund flows.

In summary, the principal objective of this dissertation is to understand how the media coverage of mutual funds impacts the purchase decisions of investors, which are manifested in mutual fund flows. The results described in this dissertation highlight the multiple effects of media coverage on investors' attention and learning about mutual funds, and ultimately on their flows into or out of the funds.

The remainder of the dissertation is organized as follows. Chapter 2 reviews the relevant literature, including insights from Psychology applicable to my research. Chapter 3 develops the key hypotheses. Chapter 4 discusses the data and methodology. Chapter 5 studies the influence of fund characteristics on media coverage. Chapter 6 describes the Heckman selection model and the accompanying empirical analysis of the impact of media coverage on mutual fund flows. Chapter 7 presents the robustness tests and Chapter 8 concludes.

Chapter 2

Literature Review

The principal focus of this dissertation is to better understand the role played by media coverage in the investment decisions of mutual fund investors. My study draws on various branches of the literatures on mutual fund performance and inflows. In this chapter, I highlight some of the relevant literature with regard to mutual fund flows and media coverage, in addition to learning models from the rational as well as the behavioral approach. I also describe certain key insights that can be derived from the Psychology literature.

FUND FLOWS AND MEDIA COVERAGE

Most previous research on the determinants of mutual fund flows (e.g., Ippolito, 1992; Chevalier and Ellison, 1997) largely focuses on the relation between measures of past performance and future flows. However, these studies implicitly assume that it is costless for investors to gather and process information on the universe of available funds. Consistent with investors having lower search costs for those funds that they have been exposed to through the media, Sirri and Tufano (1998) find that mutual funds receiving more media attention receive correspondingly higher inflows.

Sirri and Tufano (1998) primarily study the performance-flow relationship and the determinants of mutual fund flows, especially in relation to investor search costs. They classify media coverage about mutual funds as a measure of investor search costs and conduct a preliminary study on the impact of media coverage on mutual funds. They hypothesize that funds mentioned more often in the media are more easily recognized by investors and thereby resulting in lower search costs for the investors. They analyze the determinants of media coverage and find that funds with extreme performance garner a higher share of media attention. Their study provides some evidence that a larger share of current media cites is related to faster current growth. Their analysis is however restricted by their measure for media attention. A simple count of references to each fund in Lexis/Nexis does not provide much opportunity for a more detailed analysis. The database for my analysis, on the other hand, contains detailed classification for each article in terms of tone and type of media coverage. This allows for a richer exploration of the impact of media coverage on mutual fund flows.

Reuter and Zitzewitz (2004) look at mutual fund recommendations from a sample of six publications – three personal finance magazines (*Money*, *Kiplinger's Personal Finance*, and *SmartMoney*) and two newspapers (*Wall Street Journal*, and *New York Times*) along with *Consumer Reports*. They analyze the effectiveness of expert opinion and its influence on mutual fund investors. The empirical results on the impact of media mentions is restricted to the evidence that positive recommendations are associated with a significant increase in fund assets over 12 months, and that too for *New York Times* and *SmartMoney* only. Their focus is more on the relative influence of media mentions and advertising, as well as the bias in advertising. They present evidence that personal finance magazines are more likely to recommend the funds of their advertisers. They do not find this effect for the other three publications in their sample. This bias should not affect my results as the only personal finance magazine from their list that also appears in my list of 12 publications is *Money*.

These studies suggest that investors rely on the media and advertising when deciding which mutual funds to buy and sell. Del Guercio and Tkac (2002) find that Morningstar ratings influence fund flows. I find that media coverage, both positive and negative, has a strong influence on the financial decisions of mutual fund investors by facilitating in their learning about mutual funds.

INVESTOR LEARNING MODELS

Several recent models for judging mutual fund performance and for examining the dynamics of the mutual fund flow-performance relation rely on investor learning about the mutual fund managers' skills through time (Berk and Green, 2004; Lynch and Musto, 2003; Pastor and Stambaugh, 2002; Baks, Metrick, and Wachter, 2001). When investors have imperfect information about expected returns or cash flows, they must learn about the unknown price process from whatever information is available. Investor learning has been formally modeled using Bayesian analysis in most studies in the rational literature on learning models. Typically in such models, the investors (and sometimes the managers) update their posteriors on the basis of the history of observed returns as Bayesians.

Berk and Green (2004) find that there is learning about managerial ability from past returns. Investors learn about managerial ability by observing the realized excess returns the manager produces. This learning is then the source of the relationship between performance and the flow of funds. Investors and the manager update their posteriors on the basis of the history of observed returns as Bayesians.

Lynch and Musto (2003) endogenize the flow-performance relationship. Similar to Berk and Green (2004), investors and managers learn about managers' abilities and the profitability of their strategies from past returns. However in Lynch and Musto (2003), differences in ability lead to persistent differences in performance, while Berk and Green (2004) show that rational learning and strong response of flows to performance can be consistent with no persistence in performance. Baks, Metrick and Wachter (2001) and Pastor and Stambaugh (2002) both use the Bayesian approach to performance evaluation. In the former study the authors develop a flexible set of prior beliefs about managerial skill, while in the latter study the authors distinguish beliefs about the ability of the benchmarks to price passive assets from beliefs about the potential skill of fund managers.

The behavioral finance papers adopt an alternate approach that Bayesian updating is not fully descriptive of human behavior. In view of this, many recent studies have explicitly modeled how decision-making occurs in a way that reflects psychological biases. Daniel, Hirshleifer and Subrahmanyam (1998) focus on psychological evidence as a basis for assumptions about investor behavior. They theorize that news watchers rationally use fundamental news but ignore prices. Barberis, Shleifer and Vishny (1998) propose a model of investor sentiment based on how investors form beliefs, with a focus on learning about the time-series process of a performance measure such as earnings. Further, the authors refer to under-reaction evidence indicating that current good news has power in predicting positive returns in the future.

Some recent papers have examined attention and learning effects, though not in the context of media coverage of mutual funds. Barber, Odean and Zheng (2005) contend that over time, investors have become increasingly aware of and averse to mutual fund costs. They find that investors have learned more quickly to avoid salient fees like high front-end-load and commission costs than high operating expense costs. They test the hypothesis that investors have learned to avoid front-end load fees by experience. The results of their test provide direct evidence of learning. They find that experienced fund purchasers pay, on average, about half the front-end load fees of first time purchasers.

Barber and Odean (2004) propose an alternative model of decision making based on the assumption that investors buy stocks that catch their attention. They believe that this attention-based buying results from the formidable search problem that investors face when buying a stock. The authors suggest that investors solve this problem by only considering for purchase those stocks that have recently caught their attention. They test and confirm the hypothesis that individual investors are net buyers of attention-grabbing stocks, e.g., stocks in the news. They argue that firms that are in the news are more likely to catch investors' attention than those that are not as news is a primary mechanism for catching investors' attention. I apply a similar logic to the study of mutual funds in this dissertation and provide evidence that investors respond to media coverage of mutual funds by the significant increase in fund flows.

Merton (1987) advances the investor recognition hypothesis to describe the portfolio formation of informationally constrained investors. He notes that individual investors tend to hold only a few different common stocks in their portfolios. He points out that gathering information on stocks requires resources and suggests that investors conserve these resources by actively following only a few stocks. His analysis also implies that changes in investor recognition will coincide with changes in investment. Merton's paper emphasizes the differences in the breadth of investor cognizance. The analysis in this dissertation emphasizes the differences in the breadth of media coverage of mutual funds that leads to investor cognizance.

INSIGHTS FROM PSYCHOLOGY

Much human learning is aimed at developing cognitive skills on how to gain and use knowledge for future use. This section presents some insights from the cognitive theory of psychology in the context of its implications for learning and attention effects of media coverage of mutual funds. After presenting the theoretical backdrop, I describe an influential cognitive model to study the learning and attention effects of the media. I conclude each sub-section by highlighting its application to my research.

Cognitive Theory of Learning

Cognitive theories of learning in the field of psychology view information processing as constructive; that is, people do not literally encode and retrieve the information that they read or hear in the media. Rather, as they comprehend, they interpret in accordance with their prior knowledge and beliefs, as well as the context in which the message is received. Comprehension of a news story emerges through a continual interaction between the content of the program and the knowledge already in our minds. The mind is always actively thinking about what we see or hear and those thoughts become an important part of the constructive process of comprehension.

The social cognitive theory was originally developed out of the stimulus-response behaviorist psychology known as *social learning theory*. Over the years this model has increasingly stressed personal agency and cognitive processes. We learn behaviors by observing others performing those behaviors and subsequently imitating them. The relevance to media occurs when the media model becomes the source of observational learning.

There are four sub-functions for observational learning from the media. First, someone must be exposed to the media example and attend to it. Second, he or she must be capable of symbolically encoding and remembering the observed events, including both constructing the representation and cognitively and enactively rehearsing it. Third, the person must be able to translate the symbolic conceptions into appropriate action. Finally, motivational processes develop by internal or external reinforcement (reward) for performing the behavior. For example, a person's purchasing behavior could be reinforced if the behavior impressed others, or more importantly, if the person received a monetary gain as a result. Social cognitive theory was initially developed in the context of studying the effects of violent media models on behavior. Although that is still the most studied application, the theory has other applications as well, as in the modeling of sexual, prosocial, or purchasing behavior.

The first sub-function for observational learning from the media is exposure and attention to the media. My dissertation conducts an empirical study of this important aspect of learning from media coverage of mutual funds. I measure exposure to the media primarily in terms of the existence of news stories about a given mutual fund in the media that the investor is exposed to. An additional measure of media exposure is given by the circulation numbers of the various periodicals and newspapers. The theory predicts that the higher the number of news stories about a fund in a given month, the more likely it will capture the attention of investors. Thus, the number of articles about a given mutual fund in the media in a month forms a measure of potential attention to the media coverage by an individual.

Capacity Theory of Attention

Capacity theory of attention assumes that a general limit exists regarding an individual's capacity to perform mental work. It also assumes that this limited capacity can be allocated with considerable freedom among concurrent activities. Capacity theory provides a theoretical framework to study how one pays attention to objects and to acts. In this context, the terms *exert effort*, *invest capacity* and *pay attention* are often used interchangeably. To explain man's limited ability to carry out multiple activities at the same time, capacity theory assumes that the total amount of attention that can be deployed at any time is limited.

Different mental activities impose different demands on the individual's limited capacity. An easy task demands little effort, and a difficult task demands much effort. When the supply of the individual's attention capacity does not meet the demands of the task, performance falters, or fails entirely. In describing his capacity model for attention, Kahneman (1973) observes that variations of physiological arousal accompany variations of effort, suggesting that the limited capacity and arousal system must be closely related.

It is hence assumed that more capacity is available when arousal is moderately high than when arousal is low. The author also assumes that momentary capacity, attention, or effort is controlled by feedback from the execution of ongoing activities, that is, a rise in the demands of these activities causes an increase in the level of arousal, effort, and attention.

Kahneman (1973) emphasizes that attention is a scarce cognitive resource. The vast amount of information available on mutual funds places a considerable strain on the investors' capacity for attention. This opens up a potential for media coverage to play a bigger role than other sources of information about mutual funds. Media coverage can also influence what information investors are more likely to attend to by capturing their attention with certain news stories.

Limited Capacity Model

One of the most influential recent cognitive models is Annie Lang's Limited Capacity Model of media information processing (Lang, 2000). Drawing on basic cognitive psychology, Lang makes two basic assumptions: (1) people are information processors, and (2) the ability to process information is limited. Both of these assumptions are directly applicable for investor learning from media coverage, because investors have thousands of mutual funds to choose from and many investors have no training in how to select a mutual fund. When it comes to investment decisions, given the vast amount of information available and the limited resources, investors have to be selective in information processing. These information processes are sometimes automatic and sometimes controlled, that is, involving conscious volition. The three major sub-processes are encoding, storage, and retrieval, which may be done partially in parallel. Given that processing resources are limited, heavy allocation to one may result in superficial allocation to another.

One of the automatic selection mechanisms steering the encoding of information is the *orienting response*. Research suggests that this physiological and behavioral response is associated with attention and stimulus intake. When our attention is captured by something, more cognitive resources are allocated to encoding the information from that source. Two major types of stimuli activate automatic selection processes: (a) information that is relevant to the goals and needs of the individual, and (b) information that represents change or an unexpected occurrence in the environment.

Historically, in the mass communication field, the sub-process of encoding has been operationally equated to measures of exposure and attention. Many communication models and theories treat this step as a simple condition on the road to information processing, but there is nothing simple about exposure and attention in Lang's model. In the model, the process of linking newly encoded information to previously encoded information (or memories) is called storage. Some bits of information will be more thoroughly stored while other parts receive only cursory storage. Storage is affected by both individual differences and resource limitations of the human information-processing system. Retrieval, the final sub-process in this model, is the process of reactivating a stored mental representation of some aspect of the message. The more associative links there are to a piece of information (that is, the more thoroughly it has been stored), the more readily retrievable it is. The model conceptualizes retrieval as an outcome associated with learning the content of a message as well as an ongoing process during message reception.

An example applied to media coverage of mutual funds may better illustrate the concepts presented in the model. An investor reading a news article about a major mutual fund in his or her portfolio will result in concurrent retrieval of what he or she knows (about mutual funds in general, about this fund in particular, and about his or her investment in that fund) to understand the article and store this new fund information into his or her associative network. In this case, the calls for attention made by the article may actually interfere with the process of storage. In addition, the individual is likely to allocate a fair number of resources to retrieval in order to integrate new knowledge with old knowledge. The mutual fund investor is much more likely to run into a resource-limited situation than the person reading a newspaper to be entertained, since the former is purposely allocating resources to storage and concurrent retrieval in order to learn and retain the content of the article concerning the mutual fund in the portfolio.

In this dissertation, my study of the impact of media coverage on mutual fund flows draws from the theory underlying the limited capacity model. As discussed earlier, the vast amount of information available coupled with limited resources force investors to be selective in information processing while making investment decisions. Since time and cognitive resources are limited, individual investors cannot analyze all the financial data optimally. Limited attention, memory, and processing capacities force a focus on subsets of available information. Thus, the basic tenets of the limited capacity model are directly applicable to investor learning from media coverage.

The intuition developed from the psychology theory presented in this chapter contributes to my study on media coverage of mutual funds. The cognitive theory of learning and the capacity theory of attention outlined in this chapter provide a theoretical basis for understanding the human psychology of learning and attention. The limited capacity model then applies concepts based on these theories to model information processing and learning from media. The media are not only influential in telling people what to think, they are also eminently successful in telling them what to think about, thus producing attention and learning effects. Mass media hence has the ability to structure investor cognitions and to effect change among existing cognitions.

One potential manifestation of the change in investors' cognition is a corresponding change in their investing behavior, which may be studied by an analysis of mutual fund flows. I conduct various analyses of the impact of media coverage on fund flows including tests of learning and attention effects of media coverage. In the next chapter I develop the relevant hypotheses from the learning and attention effects. Later chapters of this dissertation present detailed descriptions of the specific tests for these effects along with the explanations of the results.

Chapter 3

Development of Hypotheses

In this chapter, I develop the hypotheses regarding the role played by media coverage in the investment decisions of mutual fund investors. If media coverage affects investor attention and learning about a fund, one would expect that investors would change their holdings in the funds after the news event. I test this implication by examining the effect media coverage has on investor flows into the fund, while controlling for other major factors that affect fund flows.

The theory on attention predicts that firms that are in the news are more likely to catch investors' attention than those that are not. The EntreMed example presented in Chapter 1 highlights the point that even if news is publicly available, it is not incorporated into investment decisions until and unless investors pay attention. The most basic effect of media coverage is to simply make investors aware of funds that they had not previously considered. The increased awareness of mutual funds implies that investors will view these funds as possibilities for their investment opportunity. Media coverage also reduces search costs for mutual fund investors and thus has a material impact on investor fund choices. Specifically, I test the hypothesis:

Hypothesis 1: *Mutual funds with higher media coverage have greater fund flows.*

The first hypothesis tests the overall impact of media coverage, and hence I use absolute values of fund flows as the independent variable in the regressions. Thus the first hypothesis tests whether overall media coverage, irrespective of its posture, results in greater absolute fund flows. The rest of the hypotheses take into consideration the posture of the media coverage and hence directional fund flows will be used thereafter for all the regressions.

A key influence of media is knowledge leading to investor learning. The cognitive effects of learning include concerns about what is learned as well as how much is learned. I explore this aspect of learning empirically by examining the information content of the media coverage in terms of the posture of the news story, i.e., positive or negative slant in the article with regard to the fund. Investor learning from media coverage could also impact the fund's existing investors by leading them to reevaluate their current investments in such funds, resulting in either increased investments (in the case of a positive news story) or reduced investments (in the case of a negative news story). This can result in a significant impact on fund flows considering the finding by Alexander, Jones and Nigro (1998) that in choosing which mutual funds to purchase, 42% of the investors state that they rely on financial publications like newspapers and magazines for information about the funds. I specifically test the following hypothesis:

Hypothesis 2: *News articles with positive or negative posture have a stronger impact on mutual fund flows than news articles with neutral tone or no posture.*

I test for whether fund characteristics influence the probability of media coverage, and find that this is indeed the case. I propose that fund characteristics influence more than just the probability of media coverage. There is considerable finance literature on fund flows to provide evidence that determinants of fund flows include fund characteristics such as past performance and fund size. In addition to their direct impact on fund flows, I propose that these fund characteristics also impact the effect of media coverage on fund flows. Specifically, I hypothesize that better past performance will enhance the impact of positive news stories and diminish the impact of negative news stories on fund flows. In the case of fund size, I predict a smaller impact of news on flows of larger and well known funds in contrast to the more significant role played by news for smaller and lesser known funds. I gather all these predictions under the umbrella of my next hypothesis:

Hypothesis 3: *Fund size and past performance influence the impact of media coverage on mutual fund flows.*

Learning theory suggests that investors are likely to have better knowledge of older funds that have been around for a longer time. This implies that the there is a weaker learning potential for funds that are older and have a greater cumulative number of news stories. Consistent with both attention and learning theories, I posit that media coverage of younger funds could bring them to the attention of investors, and that there may be more for investors to learn about younger funds. This leads to the hypothesis that media coverage has a greater impact for such funds than older and well known funds. Specifically, I test the following hypothesis:

Hypothesis 4: *As a fund ages and investors receive additional news about the fund, the later news has smaller effects on fund flows.*

The previous hypotheses are consistent with both the attention and the learning theories. In order to differentiate between these theories, I devise a test in which I examine flow differences between two sets of matched funds. One fund in each pair has been listed in a ranking article, where the fund is assigned a numerical ranking anywhere in the text or table section of the news story.[10] The other fund in the pair is closest to the first fund in terms of performance but has not been featured in a ranking article. Thus, the pair consists of two funds that are nearly identical in performance, but only one's performance is mentioned in the press. I posit that a statistically significant difference in flows between the two funds will provide strong evidence for the attention effects of media coverage of mutual funds. This leads to my final hypothesis:

Hypothesis 5: *A fund is more likely to catch the attention of investors when it makes a conspicuous appearance in the media, such as being portrayed with a numerical ranking.*

[10] For example, an article ranking the top three funds of the year.

Chapter 4

Data and Methodology

MUTUAL FUND DATA

I construct a sample of mutual funds that existed in the CRSP mutual fund database over the 1994-2000 time period. In order to provide tests of investor attention and learning about these funds through time, I include only funds that existed over the entire sample period. I constrain the sample to growth funds for two reasons. First, most previous research on mutual fund flows (e.g., Chevalier and Ellison, 1997; Sirri and Tufano, 1998) has focused on funds with the growth objective. Second, funds with this objective appear to hold the most investor interest and are thus the most likely to generate media coverage and reflect investor attention and learning effects, if they exist.

During the latter part of my sample period, mutual funds were sold in share classes. The same fund, i.e., the same underlying portfolio of assets, can be purchased through different types of shares with different fees (and thus, different net returns). See Appendix A for a brief description of mutual fund share classes.[11] I identify 406 share classes for the 286 growth funds in the sample. For each fund share class, I obtain information about its total net assets, returns, age, expense ratios, and load fees from the CRSP Mutual Fund database. Although share classes of the same fund differ in their

[11] See Nanda, Wang and Zheng (2004) for an analysis of the issuance of additional share classes.

flows, returns, and expense ratios, the vast majority of news articles concern the fund rather than a particular share class. Further, without controlling for the commonality among the share classes of the same fund, my standard errors will be inflated. Consequently, I combine the share class observations into a single observation for each fund for each time period. I use the total net assets of the share classes to construct a value-weighted variable for each of the fund characteristics.

Variables of Mutual Fund Characteristics

The following variables are used in this study to represent fund characteristics:

Net Fund Flow

I calculate net flows into fund *i* over period *t* with the following equation from Sirri and Tufano (1998):

$$\text{Flow}_{i,t} = \frac{\text{TNA}_{i,t} - \left\{\text{TNA}_{i,t-1} * \left(1 + \text{R}_{i,t}\right)\right\}}{\text{TNA}_{i,t-1}}$$

where $\text{TNA}_{i,t}$ represents fund *i*'s total net assets at the end of period t and $R_{i,t}$ represents fund *i*'s return over period *t*. In this measure, $\text{flow}_{i,t}$ represents the aggregate of all share classes for a particular fund and is expressed as a percentage flow.

Market Flow to Growth Funds

I construct a proxy for the average market flow to funds with the growth objective over period *t* by calculating the average flow for all funds in the sample over period *t*.

Log Lag TNA

The Log Lag TNA variable is the natural logarithm of the previous period's Total Net Assets (TNA). The monthly TNA is obtained as of the end of the specified month for which the data applies and is expressed in millions of dollars.

Fund Return for period t

The total return per share on fund *i* from time *t-1* to time *t* is calculated by the following equation:

$$R_t = \left(\frac{NAV_t}{NAV_{t-1}}\right)\left\{\prod_{j=1}^{J}\left(1+\frac{X\_AMT_j^D}{RE\_NAV_j^D}\right)\right\}\left\{\prod_{k=1}^{K}\left(\frac{X\_AMT_k^S}{RE\_NAV_k^S}\right)\right\}-1$$

where

NAV_t is the Net Asset Value at the end of the period t.

J is the number of dividend or capital gains distributions during the period.

K is the number of NAV splits during the period.

$X\_AMT_j^D$ is the j^{th} dividend or capital gains distribution during the period.

$RE\_NAV_j^D$ is the NAV at which the j^{th} dividend or capital gains distribution was reinvested.

$X\_AMT_k^S$ is the number of new shares per $RE\_NAV_k^S$ of old shares investors received in the k^{th} NAV split over the period.

$RE\_NAV_k^S$ is the number of old shares investors traded in for $X\_AMT_k^S$ new shares in the k^{th} split. That is,

$\frac{X\_AMT_k^S}{RE\_NAV_k^S}$ is the split ratio for the k^{th} NAV split.[12]

[12] Source: CRSP Survivor-bias free US mutual fund database guide.

Return Volatility for month t

For each fund, for each month, I calculate the return volatility as the standard deviation of monthly return over the previous 12 months.

Return previous year

I construct this variable as a 12-month moving statistic for each month. For each fund *i*, I calculate this variable as:

$$\text{Return previous year}_{i,t} = \left\{ \prod_{j=1}^{12} (1 + R_{i,t-j}) \right\} - 1$$

where $R_{i,t}$ is the total monthly return for fund *i* over period *t*.

Expense Ratio

Expense Ratio is the percentage of the total investment that shareholders pay for the mutual fund's operating expenses over the calendar year.

Total Load Fee

The total of all maximum front, deferred and rear-end load charges is a percentage total of loads applied to a fund.

Fund Age

I calculate fund age as the number of years since the year the fund was organized.

Young Fund Flag

Yng_i = 1 if fund *i* is in the youngest 1/3 of the sample by age, else = 0

Old Fund Flag

Old_i = 1 if fund *i* is in the oldest 1/3 of the sample by age, else = 0

Descriptive Statistics of Mutual Fund Characteristics

Table 1 provides summary statistics for each of the fund characteristic variables during my sample period. Panel A of Table 1 provides the key descriptive statistics for my full sample. Panel B provides the same statistics for each of the fund characteristics for the two sub-samples based on the existence (or lack thereof) of media coverage. Panel B also provides t-statistics for the differences in the means of the fund characteristics for sub-samples with and without media coverage. As the table shows, the growth funds in the sample vary greatly in size, ranging from about $22 million at the tenth percentile to well over $3 billion at the ninetieth percentile, with a mean (median) of $1.7 billion ($245 million). It is interesting to note that the mean fund size for observations with media coverage is over four times the mean fund size for observations with no media coverage, with the t-statistic for the difference in the two means being highly significant. This suggests that a fund's size may influence the probability of media coverage for that fund. I explore this further when examining the determinants of media coverage later in the dissertation. Figure 1 shows that the average total net assets under management increases from under $750 million in 1994 to almost $3 trillion in 2000. This is consistent with changes in mutual fund assets in general over the sample period.[13]

On average, over the sample period, the monthly flows into these funds are a little over 1% of their total net assets, as can be seen in Panel A of Table 1. The variation across the funds is large, ranging from a negative monthly flow of 2.7% at the tenth percentile to a positive monthly flow of 4.3% at the ninetieth percentile. The average

---

[13] Look at ICI website for data in general.

flow for funds in the news is distinctly greater than the average flow for funds not in the news, with the t-statistic for the difference in the means significant at the 1% level. In fact, the median fund without media coverage receives no net flows in that month. The difference between funds with and without media coverage is more obvious at the higher end (see the numbers for the 90th percentile) for both fund size and fund flow variables.

As can be seen from Figure 2, the sample funds experienced negative outflows towards the latter part of the sample period. In 1998 U.S. stock indexes experienced their largest intrayear declines since 1990 and net new cash flow to equity funds slowed to $157 billion in 1998 from $227 billion in 1997.[14] While equity funds enjoyed slightly better net new cash inflows in 1999, it was still lower than the 1997 level. Funds with objective of 'Growth & Income' saw net new cash flow drop from $84 billion in 1997 to $31 billion in 1999 and experienced a net cash outflow of $32 billion in 2000.[15] Further, the net flows across individual domestic equity funds were more concentrated than normal, with 90 percent of domestic funds experiencing a combined net outflow of $109 billion.[16] While net new cash flow to equity funds rose in 2000, the bulk of the inflow was posted in the first quarter of the year, and flows slowed markedly as the year progressed.[17] This was largely on account of domestic equity markets experiencing their sharpest correction in many years. 56 percent of all stocks listed on U.S. exchanges declined in 1999. The NASDAQ declined 40 percent in 2000, the S&P index dropped 10 percent, and the Dow Jones industrial index fell 6 percent.[18] Additionally, the smaller

[14] Source: Investment Company Institute
[15] Source: Investment Company Institute
[16] Source: Investment Company Institute
[17] Source: Investment Company Institute
[18] Source: Investment Company Institute

percentage flows to funds in the second half of the sample period is consistent with increasing assets under management as shown in Figure 1.

Figure 3 depicts the temporal changes in total monthly return for the mutual funds in my sample. The monthly returns on the funds average 1.26% over my sample period, with a range from the tenth to the ninetieth percentile between –4.14% to 6.42% per month, as shown in Panel A of Table 1. Panel B shows that there is no significant difference in the in the mean monthly return between the funds with and without media coverage. The distribution of expense ratios is substantially tighter with an average of 1.27% of total net assets. The mean expense ratio is higher for funds with media coverage than funds without media coverage, and the t-statistic for the difference is significant at the 1% level. This difference is more pronounced at the higher end, as shown by the numbers for the 90th percentile. The average across the load fees is 2.09%, with the median fund having a higher load fee of 3.91%.

The age of the funds in my sample ranges from 3 years at the tenth percentile to 41 years at the ninetieth percentile with the average age over 15 years. Panel B of Table 1 shows that the average age for funds with news articles is 22.54 years while the average age for funds without news coverage is 13.85 years, with the t-statistic for the difference in the two means being highly significant. This suggests that older funds are more likely to receive media coverage than younger funds. This influence is part of the learning effects of media coverage and I explore this further when examining the relationship between age and media coverage later in the dissertation. This difference in age seems to span the entire range of fund age, with the proportion of the fund age statistic for the two groups remaining quite steady from the 10th to the 90th percentile.

NEWS DATA

The news data is a unique database of nearly 10,000 news stories built specifically for the purpose of this study. I searched for mentions of the 286 funds in my sample in 12 major publications through the Dow Jones Retrieval Service (now called Factiva) for the 1994-2000 sample period. In the daily publications are included the highest circulation newspapers with the largest number of stories about mutual funds available from Factiva, as can be seen by the circulation numbers provided in Table 2. For the weekly or monthly publications, I include four major business publications (*Barron's, Business Week, Forbes* and *Fortune*), one national news magazine (*U.S. News and World Report*), and one personal finance magazine (*Money*), each of which had the largest number of stories about mutual funds in their category. From the 12 publications, I locate a total of 9,984 articles that mention the funds in my sample and these news articles are further classified as detailed in the following section.

Methodology of Classification of News Articles

I personally collected and classified around a third of the news articles in the database. I also trained and supervised a team of 6 research assistants to help in the collection and classification of the remaining news articles. I individually trained and monitored each member of the team for a significant period of time. After each member was fully trained, I would assign funds for which they were to search and classify news articles. I picked a random sample of 20% of the articles classified by each of them and reclassified them to ensure total compliance with the training given. Please see Appendix B for an elaboration of the training methodology. Each of the 9,984 news articles was read and subsequently classified in terms of a given set of characteristics. The classification and the corresponding characteristics of the news articles were then entered in a spreadsheet. These are detailed in Appendix B.

The classification of the news articles is primarily on two fronts: tone and type. The tone of the news articles range from Positive to Negative with Neutral in the middle. News articles that portray a fund in a clearly positive (negative) light are classified as Positive (Negative). News articles that exhibit a marginal positive (negative) posture towards the fund are classified as Neutral- Positive (Neutral- Negative). News articles that do not exhibit any posture towards the fund are classified as Neutral.

The classification of the type of news article is determined by the extent of the portrayal of the fund rather than the posture as in the case of the tone classification. The type of news article is classified in the following order. First, it is seen if the article can be classified as a Feature. If an article focuses on a fund to the extent that a considerable portion, if not all, of the article portrays the fund, then the article is classified as a Feature. If bulk of the article does not focus on the fund, but the article details an interview, usually with a fund manager, then it is classified as an Interview.

If an article does not qualify as a Feature or an Interview, then the next two types to be considered are related to the return performance of the fund. If the fund is assigned a numerical ranking anywhere in the text or table section of the news story, then the story is classified a Ranking type. If the article discusses the fund's performance but does not assign a numerical ranking, then the article is assigned Performance as the type.

If a news story does not qualify for any of the above types of classification and merely mentions the fund in a non-performance context, then the news story is classified as Mention. When the fund is mentioned nowhere in the text part of the article and only appears in a table (without a numerical ranking), then the article is classified as a Tables type.

Variables of News Data

I build on the above classification of news articles to construct the other news variables. The basic news variables are of two kinds:

- News flag or dummy variables to represent the existence of media coverage
- News count variables to represent the amount of media coverage

The two kinds of news variables provide different measures of media coverage. The news flag or dummy variable is a simple tool to study the impact of the existence of media coverage. For example, it can be employed in a regression model to test for whether mutual funds enjoy enhanced fund flows on account of being portrayed in the news media. The news count variable, on the other hand, provides a measure of the extent or amount of media coverage. The motivation for this is to capture the additional effect of the strength of media coverage by including the actual number of news stories portraying the fund during a given month. To draw a parallel, the dummy variables capture the effect for *if* media coverage exists, while the news count variables capture the effect for *how much* media coverage the fund receives.

An article flag variable represents the existence of *any* media coverage while separate flag variables represent the posture of the article towards the fund. A positive flag denotes the existence of news articles classified as Positive or Neutral-Positive while a negative flag denotes the existence of news articles classified as Negative or Neutral-Negative. The Neutral flag comprises solely of Neutral articles. The news count variables are constructed along similar lines. The total article count is the sum of all articles for fund *i* in period *t*. The positive (negative) count is the sum of the number of Positive and Neutral- Positive (Negative and Neutral-Negative) news stories for the fund in the month. The Neutral count variable is simply the number of Neutral articles for a given fund in a given month.

Additionally, there are two special count variables given by:

- Cumulative Article Count, which is the cumulative number of news articles for each fund since the start of the sample period (January 1994). The higher the cumulative article count, the more well known the fund can be said to be.

  $$\text{cumARTC}_{i,t} = \sum_{t=1}^{n} \text{ARTC}_{i,t} \text{, where}$$

  t = 1 is the first period in the sample (Jan 1994)

  t = n is the current period

- Circulation news variable, which is the circulation-weighted measure of media coverage. This variable incorporates a measure of the reach of media coverage along with the amount of media coverage, thereby forming a good tool to measure the impact of media coverage.

$$\text{Circ}_{i,t} = \text{A}_{j,i,t} * \frac{\text{C}_j}{\sum_{j=1}^{12} \text{C}_j} \qquad \text{, where}$$

$\text{A}_{j,i,t}$ = # articles in periodical j for fund i in period t

C_j = circulation of periodical j

Descriptive Statistics of News Data

Table 3 provides the frequency distribution of news articles collected and classified uniquely for this research. Panel A shows the list of publication sources as well as the number of articles from each publication. The 12 publications yield a total of 9,984 articles that at least mention the 286 funds in my sample. Of these, 62% come from daily publications and 38% come from weekly or monthly publications. In terms of individual publications, the *Wall Street Journal* carries the largest percentage of the articles regarding the sample funds (30.98%), followed by *USA Today* (14.29%), *Barron's* (13.85%) and the *Boston Globe* (10.25%).

Figure 4A presents a temporal depiction of the probability of a fund receiving media coverage. The graph does not appear to display a significant trend across time. When examined within each year, however, a quarterly trend emerges. Figure 4B charts the seasonality across months in the probability of media coverage and confirms the aforementioned trend. The trend is that the probability of a mutual fund being portrayed in a news story spikes up at the beginning of each quarter. This is likely associated with the quarterly release of mutual fund performance data. Many statistics associated with mutual fund assets and flows are released as of the end of the preceding quarter. This usually results in many stories at the beginning of each quarter that analyze the mutual fund performance in the preceding quarter, and some predictions for the quarter ahead.

Each of the 9,984 news articles in my database is classified in terms of its posture toward the fund, as described in the previous section. The results of the classifications are shown in Panels B and C of Table 3. Although the majority of articles are simply neutral, Panel B shows that an asymmetry exists in the posture of the articles with a substantially greater proportion being classified as positive or neutral-positive as compared to negative or neutral-negative. That is, over my sample period, a news article is almost thrice as likely to have mentioned a fund in a positive rather than negative light.

There are at least three potential interpretations of these results. First, it could be that more positive than negative news about growth funds occurred naturally over my sample period, which encompasses a bull market for the majority of the period. I test this by conducting a study of the impact of fund performance on media coverage. The analysis is described in detail in Chapter 5. I find that the fund's performance over the previous year has a significant impact on the posture of media coverage received by the fund. I find a significantly higher proportion of positive news stories when funds have performed well and more negative news stories when funds perform below par. These results support the hypothesis that the higher proportion of positive media coverage for my sample may be driven by a bull market for the majority of the sample period.

An alternative interpretation would be similar to the hypotheses regarding analysts' bias in forecasting earnings estimates (or recommendations), where reporters could also have positive bias in order to gain continued access to the fund's management company.[19] A third interpretation has been suggested by Reuter and Zitzewitz (2004) who maintain that some financial publications that accept advertising from mutual funds are more likely to provide positive recommendations for the advertised funds. This last possible interpretation is unlikely to explain the results seen in my study as Reuter and Zitzewitz (2004) focus on recommendations from six newspapers and personal finance magazines: *Wall Street Journal, Money, New York Times, Kiplinger's Personal Finance, Smart Money* and *Consumer Reports*. Their results derive from the personal finance magazines, of which I have only one (*Money*) out of the 12 publications covered by my sample.

Figures 5A-5C present a temporal depiction of the frequency of media coverage received by the funds in the sample. Each chart highlights a specific posture of the media coverage. Figures 6A-6C present the seasonality across months in the frequency of the funds receiving positive, neutral and negative media coverage respectively. Similar to the probability of *any* news story, these charts also display a quarterly trend. The frequency of the funds being portrayed in news stories with positive/neutral/negative posture spikes up at the beginning of each quarter. These charts likely capture the same phenomenon resulting from release of mutual fund data at the end of every quarter that is discussed earlier in this section.

In terms of types of stories, Panel C of Table 3 shows that 6.42% of the news articles are feature stories about the fund or contain an interview with the fund manager, and 7.7% are articles about fund performance or contain numerical ranking of mutual

[19] The argument is that analysts positively bias their forecasts and recommendations in order to stay on friendly terms with firm management, thus, allowing the analysts continued access to senior management. See, for example, Francis and Philbrick (1993), Das, Levine, and Sivaramakrishnan (1998), and Lim (2001).

funds, usually with regard to performance or assets. In the majority of articles (85.88%) the fund was just mentioned or included in a table. Figures 7A-7F present a temporal depiction of the frequency of media coverage received by the funds in the sample. Each chart highlights a specific type of the media coverage. Figures 8A-8F present the seasonality across months in the frequency of each type of media coverage portraying the funds in the sample. While the charts more or less portray the quarterly trend of Figures 4 and 6, Figures 8A and 8B show that more news stories are written about fund performance and contain numerical ranking of mutual funds in the first quarter than any other quarter. Figures 8E and 8F show that news articles that are feature stories or interviews with fund managers are more spread out across all the months.

Table 4 provides the descriptive statistics of news articles. Panel A shows that for any fund, having a news article is not common given that on average, less than 20% of the funds have at least one news article per month. It is even less common for a fund to have an article with a positive or negative tone. On average, 10% of the funds have positive articles in a month and only 4% have negative articles. Panel B of the table shows the frequency of the media coverage for the individual funds on a monthly basis, by the posture of the article towards the fund. On average, a fund has 0.42 article per month and half the articles do not portray the fund in a strongly positive or negative light. The median fund has no articles about it in a given month.

Panel B of Table 4 shows that the maximum number of articles for a fund in a month is 24. These articles portray Vanguard's S&P 500 Index fund. This level of media coverage occurred at two different times during the sample period, March 1997 and August 1997. The March 1997 articles focused on the large inflows to the Vanguard index fund (net inflows of $2.99 billion) and the high returns on the index, with only 8% of mutual funds earning better returns in the first quarter. The August 1997 articles were apparently a result of Fidelity Magellan's announcement that they intended to close to new investors. As the Vanguard index fund was the largest rival, much of the news centered on the expected effects to that fund. Despite such a large number of articles, 22 of the 24 articles on the Vanguard S&P 500 Index fund that occurred in August 1997 were classified as neutral articles. The finding that the Vanguard S&P 500 Index fund is the most mentioned fund in a month should not be surprising since it is one of the largest funds in the sample period.

A similar examination of the funds with the largest number of positive or negative articles in a given month suggests that return performance affects the probability of news coverage. Panel B shows that the largest number of news stories in a month that depict a fund in a positive or negative light are 12 and 9 respectively. The fund with the largest number of positive articles in any month in my sample period was Alger Funds' Capital Appreciation Portfolio, which had 12 positive articles in July 1995. These articles primarily discussed the fund's high percentage holdings in the technology sector and its consequent high performance due to the increases in prices in that sector. The fund with the largest number of negative articles in a month was Frontier Equity Fund Portfolio. The nine negative articles in October 1998 focused on this fund's position as the worst performing fund over the previous 5 years.

The preceding discussion suggests that fund characteristics, such as size and performance, seem to exert an influence on the media coverage received by the fund. The next chapter is dedicated to an in-depth analysis of this topic.

Chapter 5

Media Coverage and Fund Characteristics

In this chapter, I study the contribution of fund characteristics that influence the visibility of mutual funds. I conduct several analyses to empirically examine the impact of the fund characteristics on media coverage of the funds. I start with an economic analysis of the expected important factors that influence media coverage. I then employ univariate analyses to study the direct relationship of each factor with media coverage. Finally, I conduct a multivariate probit analysis to study the determinants of media coverage.

Evidence suggests that certain funds garner greater media coverage than others. Sirri and Tufano (1998) posit that larger funds, those with extreme (high or low) performance, and those with more volatile returns are more likely to be covered by the media. They find that, as expected, sheer media coverage is higher for larger funds and funds with more volatile returns. They also use fund complex size as a measure of search costs assuming that larger complexes are more visible and have greater brand awareness than smaller fund complexes. The individual investor is likely to be better off in a large fund that is a member of a large fund family. Previous studies have examined the effect of affiliation with large fund families to account for the economy of scale in raising fund visibility and reducing investment barriers, thereby making it easier for the fund to attract new investors.[20]

Prior work has established that mutual funds exhibit economies of scale.[21] Scale economies are exhibited in any industry when the fixed costs of running the firm can be allocated over a larger plant size or, in the case of mutual fund management groups, over more dollars of assets under management. If there are fixed costs associated with running mutual funds, then the larger funds have more assets over which those costs can be allocated. The benefit of economies of scale can then be passed along to shareholders as lower expense ratios. The increased stability of the fund also decreases expenses because of reduced transaction costs, since the fund need not buy and sell as frequently to meet redemptions. A stable and fast-growing fund may also be able to negotiate more favorable terms for its contractual relations. Assets under management can grow by attracting new money or posting strong returns. Successful funds that have received large inflows of investment are likely to find it easier to attract additional capital from investors. Thus, it is reasonable to expect that larger firms would tend to receive a greater level of media attention. I test this implication with my data and find evidence to support it. I present the analysis later in this chapter.

Mutual funds more frequently mentioned in newspapers and magazines are more likely to be well-recognized by consumers. Funds receiving media attention are also more likely to grow faster and have a stronger performance-flow relationship. A number of previous studies have established that mutual fund investors chase performance.[22] Flows

---

20 Huang et al. (2006), Dowen and Mann (2004).
21 Baumol et al. (1990), Malhotra and McLeod (1997)
22 Chevalier and Ellison (1997), Sirri and Tufano (1998)

into and out of mutual funds are seen to be strongly related to lagged measures of excess returns. Berk and Green (2004) model investors that chase performance and make full rational use of information about funds' histories in doing so.

Sirri and Tufano (1998) find that the media seem to treat good and bad performers almost equally. This may be attributed in part to their measure of media attention indicating the sheer information flow about a fund, rather than positive, negative or neutral news about the fund. Their results suggest a U-shaped relationship between performance and media attention. Extreme performance - whether high or low - gets media attention, and almost at the same rate. A fund or a stock that soars or dives catches peoples' attention. It is a matter of routine for news agencies to report the prior day's big winners and big losers. Barber and Odean (2004) find that individual investors tend to be net purchasers of stocks on high attention days, following extreme price moves. Investors are likely to notice when stocks have extreme one day returns. Such returns, whether positive or negative, will most often be associated with news about the firm. The news driving extreme performance will catch the attention of some investors, while the extreme return itself will catch the attention of others. Even in the absence of other information, extreme returns can become news themselves. *The Wall Street Journal* regularly reports the previous day's top gainers and losers.

Thus, I propose that it is reasonable to expect funds with extreme performance to attract greater media attention. I also posit that extreme high performance will garner more media coverage exhibiting the fund in a positive light, while extreme low performance will attract attention that is less flattering. I test these hypotheses and find evidence in support of them. I present the results along with a detailed analysis later in the chapter.

It follows from the above analysis that funds with higher volatility of returns will garner a higher share of media attention. Huang et al. (forthcoming JF) find that investors recognize and respond to risk levels of fund performance when allocating their wealth among funds. Sirri and Tufano (1998) find that media coverage is higher for funds with more volatile returns. Funds with extreme performance are more likely to have volatile returns, and thus more likely to attract the attention of the media. While the overall media coverage is likely to be higher for funds with volatile returns, I do not expect all media coverage to respond in the same manner to return volatility. Return volatility is widely used as a measure of risk, and risk is not regarded as a desirable characteristic when considering an investment. Hence, I propose that return volatility is likely to result in more news articles with a negative portrayal of the fund than a positive slant. I test these hypotheses and find that return volatility does increase the propensity for a fund to receive media coverage. In particular, volatility plays a strong and significant role for negative media coverage.

I thus identify the three fund characteristics that are likely to be key determinants of media coverage of mutual funds. In the next step, I study the direct relationship of each of these factors with media coverage, and then proceed to a multivariate analysis of the determinants of media coverage.

UNIVARIATE ANALYSIS

Media Coverage and Fund Size

As highlighted by the discussion above, one would expect larger funds to be more heavily followed as they have more shareholders, implying that a larger proportion of the publications' readers could be that fund's shareholders and consequently have an interest in the news story. Table 5 shows that there is a 42% correlation between a fund's total net asset value and the average number of total news articles about the fund per month.

To determine whether fund size influences media coverage, I divide the funds into size quintiles after sorting by the fund's total net assets. For each size quintile, I then calculate the average number of articles per month. As Panel A of Table 6 shows, being a large fund greatly increases the probability of media coverage. The largest funds have an average of more than one article per month, which is around four times as many articles as the other quintiles. The impact of fund size on the amount of media coverage received by the fund is most striking in the case of articles with a positive tone. Funds in the largest size quintile receive almost six times the amount of positive media coverage than funds in the smallest size quintile.

In contrast, the relation between fund size and negative articles is in the opposite direction. Funds in the smallest size quintile have the largest number of negative articles, and this is twice the average number of negative articles for the funds in the largest size quintile. The table further shows that being in the largest size quintile also helps in garnering the lion's share of neutral news articles as well. This is confirmed by the example from the previous chapter. The maximum number of neutral articles in any month was 22 and this occurred in August 1997 for Vanguard S&P 500 Index fund, one of the largest funds in my sample.

Media Coverage and Fund Performance

To determine whether fund performance influences media coverage, I divide the funds into performance quintiles after sorting by the fund's total return for the previous year. For each performance quintile, I then calculate the average number of articles per month. As Panel B of Table 6 shows, being a good performer greatly increases the probability of media coverage. Again, the impact of fund performance on the amount of media coverage received by the fund is most striking in the case of articles with a positive tone. Funds in the highest performance quintile receive almost four times the amount of positive media coverage than funds in the lowest performance quintile, and nearly twice that of the other quintiles.

Looking at the other side of the coin, being a poor performer draws its own share of publicity, though hardly flattering. Funds in the lowest performance quintile have the largest number of negative articles, and this is more than thrice the average number of negative articles for the funds in the highest performance quintile. This suggests that fund performance plays a key role in determining the posture of the media coverage received by the fund. In contrast to the case of fund size, fund performance does not seem to have a strong impact on media coverage that has no posture. The average number of neutral articles is marginally higher for the best performers and does not vary significantly across the other quintiles.

I conduct a study to specifically test the impact of fund performance on the posture of media coverage. I divide my sample into positive and negative bins based on whether past fund performance was positive or negative. Then for each bin, I obtain the proportion of news stories that are positive and the proportion of news stories that are negative. I repeat this process for two horizons of past performance: previous month's return and previous year's return. The results of this analysis are provided in Panel A of Table 7. For the positive bins, the proportion of positive news articles is stable at around 40% across both horizons of fund return, as is the proportion of negative news articles at the 10% level. The negative bins display a different pattern. For a shorter horizon, the proportion of positive news articles still dominates the proportion of negative news articles. However, the negative bin for previous year's return displays an inverse pattern – the proportion of negative news articles soars to 40% while the proportion of positive news articles drops to 11%. This is consistent with the results from the multivariate analysis discussed later in the chapter, where I show that the previous year's return has a significant impact on the probability of a fund receiving positive or negative media coverage.

I posit that the anomaly of higher proportion of positive news stories for the negative bin of previous month's return can be attributed to the combined influence of two factors: the weaker influence of previous month's return in comparison to previous year's return, and the powerful influence of fund size on positive media coverage. To test this hypothesis, I obtain the smaller half of the sample by restricting the fund total net assets (TNA) to below the median fund TNA. I then repeat the process of separating funds into positive and negative bins based on past performance, and then obtaining the proportions of positive and negative news stories. The results of this test are provided in Panel B of Table 7. I find evidence in support of my hypothesis. By restricting the sample to smaller funds, I exclude the phenomenon associated with large funds attracting more positive media coverage. I now find that the negative bin for previous month's return has a higher proportion of negative news stories than positive news stories. Further, I find that the disparity in the positive versus negative posture of media coverage is weaker for the shorter horizon of past performance. This is consistent with the hypothesis of previous month's return exerting a weaker influence on the posture of media coverage in comparison to previous year's return, which will be further confirmed by the multivariate analysis discussed later in the chapter.

A comparison of the two panels of Table 7 yields another insight into the interplay between fund size and performance with regard to media coverage. Positive performance over any horizon is rewarded by positive media coverage, irrespective of fund size. The media treats small firms equally well, as long as they have performed well. When it comes to negative performance however, small funds bear the brunt of the media criticism. Comparing the negative bins for the previous month's return, I find that small funds are penalized for poor performance while larger funds do not seem to be hurt by their poor returns. Comparing the negative bins for the previous year's return, I find that small funds are penalized much more severely than the average fund in the sample.

Media Coverage and Return Volatility

Table 5 shows that there is a 4.8% correlation between a fund's return volatility and the average number of positive news articles about the fund per month, and a correlation of 14.9% correlation between a fund's return volatility and the average number of negative news articles about the fund per month. This suggests that volatility of fund performance may play a significant role in influencing the media coverage of mutual funds. To study this relationship, I divide the funds into volatility quintiles after sorting by the volatility of the fund's return for the previous year. For each volatility quintile, I then calculate the average number of articles per month. The results are provided in Panel C of Table 6. Consistent with the hypothesis that overall media coverage is likely to be higher for funds with volatile returns, I find that the average number of news stories for the highest volatility quintile is nearly twice the average number for the other quintiles. The highest volatility quintile draws greater numbers of positive and neutral news stories as well. Consistent with the concept of return volatility as a measure of risk, funds in the highest volatility quintile received thrice as many negative news articles as the average of funds in all other volatility quintiles. The stronger impact of return volatility on negative media coverage is further confirmed by the multivariate analysis discussed in the following section.

The above univariate analyses identify fund size, performance and volatility as fund characteristics that seem to impact the probability of the fund receiving media coverage. I now turn to multivariate analyses to further explore the relationship between fund characteristics and media coverage of mutual funds.

MULTIVARIATE ANALYSIS

I run several probit regressions to examine the fund characteristics that affect media coverage of mutual funds. I build on the results from the univariate analyses and include fund size, past return performance measures, and return volatility as independent variables in these regressions. I also construct a *Small* dummy variable that equals one if the fund's size (as measured by its total net assets) is smaller than the median fund size of the sample, and zero otherwise. This dummy variable is interacted with the fund return for previous year. I run the regressions using different categories of media coverage as the dependent variable. Table 8 provides the key results from these tests. In Model 1, I examine the influence of the fund characteristics on the probability of any news story. That is, the dependent variable is 1 if the fund had any news article appearing for that month and 0 otherwise. In Models 2 and 3, I examine the probability of news coverage according to the tone of the coverage, with the focus in Model 2 being on the probability of a positive news story and the focus in Model 3 being on the probability of a negative news story.

The results across all three models of Table 8 show that fund size is an important determinant of media coverage. In the case of Models 1 and 2, the probability of a news story appearing about a particular fund increases in the size of the fund. That is, larger funds have a higher likelihood of receiving any kind of media coverage and there is a higher likelihood that the media coverage will be in a positive light. In the case of Model 3, the probability of a negative news story appearing about a particular fund decreases in the size of the fund. These results are consistent with the results from Table 6 discussed earlier. The results also indicate that the impact of fund size seems to be strongest for positive news stories, followed closely by the overall media coverage. Fund size does not seem to have as much of an impact on negative news stories.

I include return performance over two horizons to fully capture the influence of fund performance on media coverage. Previous month's fund return influences the probability of overall media coverage, but does not seem to play a significant role when it comes to the posture of the news stories. Return performance over the previous year seems to have a stronger effect, in terms of economic and statistical significance, on the probability of a positive or negative news story in comparison to the probability of a news mention in general. Consistent with results from Table 7, the probability of a positive news story is increasing in the previous year's return while the probability of a negative news story is decreasing in the previous year's return. This suggests that news articles that show a positive or negative posture with regard to a fund tend to take into consideration a fund's performance over a longer horizon before taking a stance on the fund's performance.

The interaction term between the *Small* dummy and the previous year's return is strongly significant across all three models. In Model 1, being a small fund weakens the probability of overall media coverage, negating the weakly positive impact of previous year's return. In the case of Model 2, the interaction term is positive. This implies that the incremental impact of better performance on the probability of positive media coverage is higher for a small fund. It is thus more critical for a small fund to perform well in order to receive positive media coverage. The strongly negative coefficient for the interaction term in Model 3 indicates that it is extremely costly for a small fund to have performed badly for a year. Being small and having a bad year of performance are individually damaging factors in terms of attracting negative media coverage. The combination of the two factors, as signified by the interaction term, is lethal for a fund's media coverage. This inference is consistent with the results from Table 7.

The volatility of returns over the previous year is seen to have a strong positive influence on the probability of overall media coverage. The impact is slightly weaker for news stories with a positive posture, suggesting that return volatility is less desirable for a complimentary portrayal by the media as compared to an average news story. The much stronger impact of return volatility seen in Model 3 for negative news articles is indicative of the disfavor with which highly volatile fund returns are regarded by the media and mutual fund investors. This is consistent with return volatility being widely used as a measure of risk when evaluating an investment opportunity.

This chapter has focused on the relationship between media coverage and mutual fund characteristics. The next chapter is dedicated to an in-depth analysis of the impact of media coverage on mutual fund flows.

Chapter 6

Media Coverage and Fund Flows

The results from the previous chapter provide evidence for the impact of mutual fund characteristics on media coverage of the fund. It is thus highly unlikely that all mutual funds have an equal probability of receiving media coverage. Previous research on determinants of fund flows has highlighted the influence of mutual fund characteristics on fund flows.[23] Selection bias may occur when funds that are likely to receive media coverage also have a greater propensity to receive higher inflows. Then the positive effects of media coverage on mutual fund flows may be overstated. Propensity for attracting flows is in the error term of the fund flow regression and is correlated with media coverage.

There are specific problems that arise when estimating a regression model with samples that may not be random. If the determination of which values of the dependent variable of a regression are to be observed is related to the unobservable error term in the regression, then methods such as ordinary least squares (OLS) are in general inappropriate. The key concern with non-random selection is that the inference based on the observed group may not extend to the unobserved group.

[23] Ippolito(1992), Gruber(1996), Chevalier and Ellison(1997), Sirri and Tufano(1998), Del Guercio and Tkac (2002)

Selection bias can be thought of as a form of omitted variable bias.[24] I will elaborate further on this while detailing the sample selection model later in the chapter. As the residual captures the effects of all omitted and imperfectly measured variables, any regressors that are correlated with the unmeasured or mismeasured factors will end up proxying for them. This is an issue because if a regressor ends up proxying for any important factors that are left out, one cannot interpret its estimated coefficient as the effect of that regressor per se, since it also captures part of the effect of the omitted or mismeasured variables.

The selection bias problem is usually addressed by employing a sample selection model. In sample selection models, one or several dependent variables are observed when another variable takes certain values. The basic idea behind these models is to estimate the following pair of regressions. The first is a probit regression predicting the probability of selection. The second is usually a linear regression for the outcome of interest as a function of the selection variable, controlling for observable confounders.

The selection bias arises if the (unobservable part of the) criterion to select into the sample is correlated with the (unobservable part of the) outcome of interest. Ideally there would be some explanatory variables in the selection regression that do not belong in the outcome regression. Strictly speaking, this is not necessary, but it helps to identify the effect of the selection on the outcome and makes the estimates more robust.

As far as testing for selection bias is concerned, the diagnosis is also the cure. One can obtain an estimate of rho (ρ), the correlation between the error terms of the two equations, sigma (σ), the standard error of the outcome regression and lambda ($\lambda = \rho*\sigma$).

---

[24] Heckman (1979)

If ρ is positive (negative), then the estimated effect of selection from single-equation estimation will generally be biased away from zero (towards zero). If $\rho=0$ (or equivalently, if $\lambda=0$, since $\sigma>0$), there is no selection bias and one can present the single-equation estimates. If $\rho\neq0$, there is selection bias and one should present the estimates from the selection model instead.

THE HECKMAN MODEL

There are two approaches to the Heckman model, introduced by James J. Heckman in 1974 and 1976 respectively. Heckman (1974) derives the sample likelihood function for the model and presents estimates obtained from optimizing the likelihood function. Later, Heckman (1976) proposes a simple two-stage estimator that permits estimation of the model by least squares and probit analysis. He provides an empirical example to show that the estimator yields estimates close to those obtained from the previous maximum likelihood estimation. Heckman (1979) clarifies and extends the analysis in the previous (1976) paper and derives the asymptotic distribution of the simple estimator. The original models have subsequently been generalized, by Heckman and by others. In Appendix C, I describe the derivation of both approaches to the Heckman model in detail. A brief description of the Heckman model is provided here.

The Heckman model can be defined by a set of two equations that describe the selection regression model (probit) and the outcome regression model (linear). The basic idea of the model is that the outcome variable is only observed if some criterion, defined with respect to the selection variable, is met. The selection model characterizes the propensity to be included in the sample. In the context of this dissertation, this is the propensity for a fund to receive media coverage. The outcome model estimates the expected value of fund flows (the variable of interest), conditional on being included in the sample and controlling for other major factors that affect flows.

Heckman's two-step procedure consists of first estimating the selection equation as a probit model. Using the probit results, the estimate of the inverse Mill's ratio or Hazard ratio (symbolized by λ) is computed for the sub sample for which the outcome variable is observed. Then, for this same sub sample, Ordinary Least Squares (OLS) model is used to regress fund flows on other major factors that affect flows, with the estimate of λ as an additional explanatory variable. The two-step model explicitly addresses bias caused by correlation of the regressor with omitted variables, by adding a term to the least squares regression that represents the non-zero expectation of the error term.

I present the derivation of the Heckman model in Appendix C and show that it is not the fact that observations on fund flows are only available for a selected sample of funds with media coverage that causes the difficulty in using the simple OLS approach. Rather, it is the fact that this selection of having media coverage is not random with respect to fund flows. The maximum likelihood approach requires the log-likelihood function to be specified. Then nonlinear optimization is used to maximize the likelihood function and calculate the standard errors of the estimates. The maximum likelihood estimates have the desirable properties of being consistent and asymptotically efficient. Therefore, I present results from the Heckman maximum likelihood model for the empirical analysis of the media coverage of mutual funds in this chapter. I present the corresponding results from the Heckman two-step model in Appendix D. This provides a robustness test to ensure that the empirical results presented in this chapter are robust to the alternate approach of the Heckman selection model.

EMPIRICAL ANALYSIS

If media coverage affects investor attention and learning about a fund, one would expect investors to change their fund holdings after the news event. I test this implication by examining whether and what kind of effect media coverage has on investor flows into the fund, while controlling for other major factors that affect flows. Given past evidence on determinants of fund flows (e.g., Ippolito, 1992; Gruber, 1996; Chevalier and Ellison, 1997; Sirri and Tufano, 1998; Del Guercio and Tkac, 2002), I control for several factors, including lag flows, fund size, past performance and fund expenses. I also control for variation across time in the overall size of investor flows into mutual funds with a variable for aggregate market flow. The variable for aggregate market flows captures the mean investor flows into all the funds in my sample during the given period. Following my earlier analysis on the relationship between fund characteristics and media coverage, my selection model is similar to the probit analysis conducted in Chapter 5. I include the fund's size, past performance measures and return volatility as determinants of the probability of media coverage in the selection regression. The definition for all the variables included in the regressions are presented in Chapter 4.

My outcome regression model to examine the relation between the fund flows and news variables during the period is then as follows:

$Flow_{i,t}$ = f (Aggregate market $flows_t$, $Flow_{i,t-1}$, Log $TNA_{i,t-1}$, Measures of past $performance_i$, Expense ratio $_{i,t-1}$, Fund $Age_{i,t}$, News $Variables_{i,t}$).

My selection regression model to predict the probability of selection is then as follows:

Probability of media $coverage_{i,t}$ = f (Log $TNA_{i,t-1}$, Measures of past $performance_i$, Return $volatility_i$)

In this section, I state the hypotheses developed in Chapter 3, describe the setup of the tests for each hypothesis, and discuss the results for each test with reference to the respective tables. I have presented the results for the probit model in Table 8, and hence do not repeat those results for each subsequent table. I present the coefficients and p-values for the variables in the outcome regression of interest for the Heckman maximum likelihood model.

The depth and breadth of media coverage

The theory on attention predicts that firms that are in the news are more likely to catch investors' attention than those that are not. The EntreMed example presented in Chapter 1 highlights the point that even if news is publicly available, it is not incorporated into investment decisions until and unless investors pay attention. The most basic effect of media coverage is to simply make investors aware of funds that they had not previously considered. The increased awareness of mutual funds implies that investors will view these funds as possibilities for their investment opportunity. Media coverage also reduces search costs for mutual fund investors and thus has a material impact on investor fund choices.

The first hypothesis relates to the key premise for the research presented in this dissertation. It tests the basic question of whether media coverage influences fund flows while the other hypotheses cover the nature of this influence. Basically, the hypothesis states that:

Hypothesis 1: *Mutual funds with higher media coverage have greater fund flows.*

This hypothesis tests the overall impact of media coverage, and hence I use absolute values of fund flows as the independent variable in the regressions. Thus the first hypothesis tests whether overall media coverage, irrespective of its posture, results in greater absolute fund flows. The rest of the hypotheses take into consideration the posture of the media coverage and hence I use directional fund flows for the other regressions.

I test the above hypothesis using two different measures of media coverage received by the fund. The first test uses the actual numbers of news stories portraying the fund each month, as a measure of the depth or strength of media coverage received by the fund. The results, presented in Model 1 of Table 9, show that having media coverage during a month has a significantly positive effect on fund flows. Monthly fund flows are increasing in the number of news articles. That is, funds that are portrayed in more news articles in a given month enjoy greater fund flows in absolute terms.[25]

An implication that can be derived from the attention theory is that the more investors that are reached by the media outlet, the greater should be the effects on funds

[25] To allow for the possibility of delay in investor reaction to news stories about the funds, I run the same regressions on a quarterly basis. I find the results very similar to the monthly results presented here.

flows. Thus, I next examine whether coverage by media outlets with greater circulation has a greater impact on investor flows. This constitutes the second test for Hypothesis 1, using circulation-weighted numbers of news stories as a measure of the breadth or the reach of media coverage. To do this, I run regressions similar to those in Model 1 in which the simple news count variable now is replaced with a weighted-circulation news count variable. This variable takes the news count variable and weights it by the proportion of the circulation of each periodical to the total circulation of the twelve print periodicals.

The results of this regression are provided in Model 2 of Table 9. The coefficients on the circulation-weighted news count variable show that there exists a stronger attention effect from having articles in the higher circulation periodicals. Comparing he two models of Table 9, I find that the coefficient for the circulation-weighted news count variable is higher than the coefficient for the plain news count variable by a factor of ten. Consistent with the attention theory, I find that the larger the reach of the mode of attracting attention, the greater the potential for influence thereafter. These results are also consistent with the example provided by Huberman and Regev (2001), in which the same article attracted much greater attention and resulted in a much stronger impact on stock prices when reprinted in the *New York Times* than when originally printed in *Nature* Magazine.

The good, the bad and the neutral

A key influence of media is knowledge leading to investor learning. The cognitive effects of learning include concerns about what is learned as well as how much is learned.

I explore this aspect of learning empirically by examining the information content of the media coverage in terms of the posture of the news story, i.e., positive or negative slant in the article with regard to the fund. Investor learning from media coverage could also impact the fund's existing investors by leading them to reevaluate their current investments in such funds, resulting in either increased investments (in the case of a positive news story) or reduced investments (in the case of a negative news story). This can result in a significant impact on fund flows considering the finding by Alexander, Jones and Nigro (1998) that in choosing which mutual funds to purchase, 42% of the investors state that they rely on financial publications like newspapers and magazines for information about the funds. I specifically test the following hypothesis:

Hypothesis 2: *News articles with positive or negative posture have a stronger impact on mutual fund flows than news articles with neutral tone or no posture.*

The results in Table 10 show that aggregate market flows, fund age, and past performance are important factors in the magnitude of the percentage flows into a fund. The greater the flows to funds in general, the greater the flows to the individual funds. The older funds have lower percentage flows as compared to newer funds. Consistent with the previous evidence regarding the influence of past performance on fund flows (e.g., Del Guercio and Tkac, 2002), I find that the fund's return performance in the previous year has a significantly positive effect on the current period's flow. I also find that flows are negatively related to the level of the expense ratio, suggesting that investors explicitly include fees in their assessments of funds.

Model 1 includes dummy variables for whether the fund had news or not in a given period, with categorization of that news into positive, neutral or negative news. The dummy variable for whether the fund has positive media coverage takes on the value of one if the fund had at least one positive news story in a given month, and a value of zero if the fund had no positive news story that month. Similar criteria apply to the other dummy variables for neutral and negative media coverage. The results for Model 1 show that having a positive news article in a period is associated with a significantly positive 1.5% boost in fund flows, while having a negative news article results in a significantly negative 1% impact on fund flows. The significance of these numbers is further highlighted by the consideration that the average fund flow for my sample is 1.08%.[26] This indicates that media coverage can have a very significant impact on mutual fund flows.

Although the news articles could simply be informing investors about the fund's recent return performance, since I control for performance in these regressions, the implication is that the news articles have an added effect.[27] Further, even just a neutral mention of a fund can have a significant effect on flows, though weaker than the stories with posture, suggesting that the appearance of a fund in a news article may increase investor awareness, consistent with the Merton (1987) investor recognition hypothesis and the attention hypothesis. These results suggest that investors learn about funds from media coverage. This is consistent with the finding in Alexander, Jones and Nigro (1998) that investors who rely on financial publications score higher on quizzes regarding their financial literacy.

---

[26] The average fund flow for the smaller half of the sample, obtained by restricting the fund total net assets (TNA) to below the sample median, is 1.46%. Even this higher average flow is considered, the impact of media coverage is still significant.

[27] I test for this effect later in Table12.

The results in Model 1 of Table 10 are based on indicator variables for whether the fund had news or not in a given period, with categorization of that news into positive, neutral or negative news. An additional issue is whether the frequency of the news articles has added effects on fund flows. To address this, I run regressions similar to those in Model 1, but after replacing the indicator variables for the existence of news with count variables for the number of news articles. The motivation for this is to capture the additional effect of the strength of media coverage by including the actual number of news stories portraying the fund during a given month. To draw a parallel, the indicator variables capture the effect for *if* media coverage exists, while the news count variables capture the effect for *how much* media coverage the fund receives. In Model 2 of Table 10, I provide separate counts for the number of positive, neutral or negative news articles portraying a fund in a period.

The results in Model 2 are generally consistent with the results in Model 1. In the case of positive and neutral news stories, I find that besides increasing in the existence of media coverage, fund flows are increasing in the number of news articles as well. That is, the more positive news articles about a fund in a given month, the greater the flows into the fund. The relationship between fund flows and media coverage is inverse in the case of negative news stories, with coefficient being equal in magnitude and significance. Fund flows are increasing in the number of neutral articles, but on a weaker note in comparison to the positive media coverage. This is particularly noteworthy given than neutral news stories comprise more than 50% of the total media coverage in my database, and yet have a much weaker influence on fund flows in comparison to positive and negative news stories. The results in Table 10 strongly support Hypothesis 2. Media coverage that contains more information, measured by posture of the news story with regard to the fund portrayed, seems to have a greater impact on fund flows than media coverage with less information.

Fund characteristics interact with media coverage

I have shown in Chapter 5 that fund characteristics, such as fund size and past performance, influence the probability of media coverage. I now explore the possibility that in addition to their direct impact on fund flows, these fund characteristics also influence the effect of media coverage on fund flows. There is considerable finance literature on fund flows that provide evidence that determinants of fund flows include fund characteristics such as past performance and fund size. In addition to their direct impact on fund flows, I propose that these fund characteristics also impact the effect of media coverage on fund flows.

I posit that better past performance will enhance the impact of positive news stories and diminish the impact of negative news stories on fund flow. Investors generally have a positive impression about a fund that has performed well. When these investors then come across media coverage portraying the fund in a positive light, it reinforces their positive impression about the fund. Hence, investors are likely to have a stronger response to a positive news story about a fund that has performed well, than a positive news story about a fund that has not done so well. In the case of fund size, I predict a smaller impact of news on flows of larger and well known funds in contrast to the more significant role played by news for smaller and lesser known funds. I have previously shown in Tables 6 and 8 that larger funds enjoy greater media coverage than smaller funds. Similar to the case of older funds, investors will have greater knowledge of large funds that have received considerable media coverage and hence the incremental impact of fresh news is likely to be smaller. These implications are captured formally in the hypothesis:

Hypothesis 3: *Fund size and past performance influence the impact of media coverage on mutual fund flows.*

In Model 1 of Table 11, I interact the natural logarithm of total net assets under fund management for the previous period with the number of news articles in the current period. The results show that in spite of controlling for size of the fund, the interaction term is significant and has a negative coefficient. The smaller percentage flows resulting from news for the larger funds indicates that media coverage plays a more significant role for smaller and lesser known funds. I do a similar analysis for Model 2 of Table 11 where I provide separate counts by posture of media coverage in terms of the number of positive, neutral or negative news articles portraying a fund in a period. The interaction terms are negative and significant for positive and neutral media coverage, thereby implying that larger funds will have lesser percentage flows from positive and neutral media coverage than smaller funds. The coefficient for negative media coverage is not significant implying that larger funds may not be as significantly impacted by negative media coverage.

To test whether fund return exacerbates or mitigates the effects of media coverage on fund flows, I add interaction terms between the counts of news articles and the fund's return performance over the previous year. In Model 1 of Table 12, I include the total number of news articles for a fund in a period and find that the interaction term is significant for the number of news stories in general. The return and news variables on their own have positive signs while the interaction term of these two variables has a negative sign. The implication of this is clarified by the results in Model 2 of Table 12, where I include interaction terms between past performance and news variables with different postures towards the fund portrayed. I find that a fund's return performance enhances the positive influence of positive media coverage. That is, funds with positive return performance and positive media coverage have greater inflows, ceteris paribus, than funds with only positive return performance or only positive media coverage. On the other hand, the interaction between return performance and media coverage for negative news stories is negative. These results are consistent with the earlier discussion in Chapter 5, particularly with the large negative coefficient for previous year's return in the probit model for negative media coverage presented in Table 8.

Media Coverage and Fund Age

An implication that can be derived from the learning models is that as a fund ages and investors receive additional news about the fund, there are smaller effects from the news, on average. There being greater learning potential for investors with regard to younger mutual funds, I hypothesize that media coverage has a greater impact for such funds than older well known funds. Specifically, I test the following hypothesis: **Hypothesis 4:** *As a fund ages and investors receive additional news about the fund, the later news has smaller effects on fund flows.*

Learning theory suggests that investors are likely to have better knowledge of older funds that have been around for a longer time. This implies that the impact of media coverage on flows is weaker for funds that are older and have a greater cumulative number of news stories. To test this implication, I provide the results of two sets of regressions that focus on the age of the fund in Table 13.

In Model 1 of Table 13, I interact the age of a fund with the number of news articles in the current period. I also interact fund age with the cumulative number of news articles about the fund beginning in 1994, the start of my sample period. The results show that both the age and cumulative number of news articles variables have significant inverse relationships with fund flows, thereby lending direct support to Hypothesis 4. That is, older funds do not receive as high an inflow as do younger funds, even controlling for the differences in sizes between the funds. I find that the interaction term between the fund age and number of news articles is negative and significant. This result is consistent with the hypothesis that investors have knowledge of older funds that have been around for a longer time and hence the incremental impact of later news is smaller.

For Model 2 of Table 13, in each period I divide the sample of funds into 3 approximately equal-sized groups according to their age: young funds that have existed 6 years or less, mid-age funds that have lives from 7 to 15 years, and older funds, with more than 15 years of existence. I construct dummy variables for the youngest and oldest groups in the regressions, omitting the middle group. I include the interaction terms between each dummy variable and the media coverage variables, namely, the current number of news articles and the cumulative number of news articles. The results show that the small, systematic differential in flows based on fund age in the effects of a news story is driven by the younger funds. The most significant impact occurs with the younger funds, which experience a 1.2% increase in flows from a news story. The news impact on the older funds is not significantly different from zero. The results in Table 13 are consistent with both the attention and the learning theories in that the difference in flows due to news coverage between younger and older funds could be a result of the younger funds coming to the attention of investors, and there may be more for investors to learn about younger funds.

Ranking articles: Test of Attention versus Learning Effects

An implication of the attention theory is that conspicuous news events are more likely to wield a greater influence than the mundane, due to their increased visibility and ability to attract the attention of a larger population. Klibanoff et al. (1998) find that the price of a closed-end country fund reacts more strongly to news about its fundamentals when the country whose stocks the fund holds appears on the front page of the newspaper. Specifically in terms of fund performance, probably the most conspicuous media coverage is when a fund is ranked as a top performer in a news article. Such news stories, where the fund is assigned a numerical ranking anywhere in the text or table section of the news story, are classified as *Ranking* articles.[28] I employ the positive ranking articles in my database to construct a test for the attention effects of media coverage.

[28] For example, an article ranking the top three funds of the year.

The previous hypotheses are consistent with both the attention and the learning theories. In order to differentiate between these theories, I devise a test in which I examine flow differences between two sets of matched funds. Each pair consists of two funds that are nearly identical in performance, but only one's performance is mentioned in the press. I posit that a statistically significant difference in flows between the two funds will provide strong evidence for the attention effects of media coverage of mutual funds. This leads to my final hypothesis:

Hypothesis 5: *A fund is more likely to catch the attention of investors when it makes a conspicuous appearance in the media, such as being portrayed with a numerical ranking.*

I conduct a test in which I examine flow differences between two sets of matched funds. One fund in each pair is listed in a ranking article, while the other is not. I then match the fund that is ranked, with another fund in my sample that is not listed in the ranking article. I match based on finding a non-ranked fund with the closest return, as long as its return for the previous year is within 50 basis points (0.005) of the ranked fund. Thus, I have two funds that are nearly identical in performance, but only one's performance is mentioned conspicuously in the press.

The results of this test are shown in Table 14. The average annual returns of the funds are 15.15% for the ranked funds and 15.01% for the non-ranked funds, a difference that is not significantly different from zero at any probability level given the t-statistic of 0.07. Yet the funds that are mentioned in the ranking articles have inflows in the month of the article averaging 3.4%, while the unmentioned funds have inflows averaging 0.27%. In a t-test of the difference between the two means, taking into account the differences in variances between the two groups of funds, the t-statistic is 3.61. This test provides strong evidence of the attention effect of media coverage. It also supports the hypothesis that attention effects are important in investors' selection of mutual funds.

In summary, I examine an underlying tenet of attention and learning models of mutual fund investors by examining sources of attention and learning, namely, news stories about funds. Specifically I examine whether media coverage of a mutual fund appears to affect investor learning and whether that learning is then manifested through changes in investment in the fund. I find that media coverage of mutual funds has a significant effect on investor flows to the fund. Further, it is evident that the media coverage affects investor learning in that not all news coverage is beneficial for fund flows. Having a negative news article on average reduces net fund flows. These results are consistent with the hypothesis that investors learn about funds through media coverage and that this knowledge affects their investment behavior. Finally, I conduct a test to differentiate between the attention and learning effects of media coverage and find strong evidence supporting the key role played by attention effects of media coverage of mutual funds.

Chapter 7

Robustness

I present some robustness tests to determine the sensitivity of the empirical results to the specification of the Heckman selection model. I consider alternate specifications of the regressions presented in the previous chapter. I employ the piecewise linear specification for the fund flow-performance relation. I also separate out the index funds in my sample to test for any difference in the influence of media coverage for these funds. I run separate regressions for the sub samples of index and non-index funds. I also run separate regressions for news stories appearing in *The Wall Street Journal,* and news stories appearing in all other publications in order to test for publication-specific phenomena. Finally, I present results from Granger causality tests run on vector auto regression models of past fund flows and past media coverage.

In Chapter 6, I employ the Heckman maximum likelihood model to account for the potential endogeneity in the relationship between media coverage of mutual funds and mutual fund flows. I also run corresponding regressions employing the Heckman two-step model to ensure that the empirical results are robust to the alternate approach of the Heckman selection model. I present these results in Appendix D. I conduct an additional robustness check on the results using another alternate specification of the regression model. I find that the economic and statistic significance of the news variables are robust for different model specifications.

Previous studies have suggested a nonlinear specification for the fund flow-performance relation.[29] Hence, I adopt the Sirri and Tufano (1998) piecewise linear specification for return performance to conduct the robustness test. I employ the piecewise linear specification using cross-sectional regressions on a monthly basis and assuming that the kinks are identical across the months. The results from the cross-sectional regressions for each month are then subjected to the Fama-Macbeth (1973) technique to aggregate the coefficients across the 84 months of my sample period. The advantage of using the Fama-Macbeth method for a pooled time-series data is that it controls for seasonality in media coverage. The seasonality of news stories in my dataset is depicted in Figures 4, 6 and 8.

The results for the piecewise linear specification are presented in Tables 15 and 16. The models in Table 15 portray overall media coverage while the regressions in Table 16 allow for the different postures of the media coverage. Table 15 shows that both the indicator variable in Model 1 and the news count variable in Model 2 have significant impact on mutual fund flows. Consistent with the empirical results presented in Chapter 6, the results in Table 16 display the strong positive impact of positive media coverage and the strong negative impact of negative media coverage. The neutral news stories with no specific posture towards the fund also have a positive impact of fund flows, as seen in the earlier results. This confirms that the results for the empirical analysis of the fund flow-media coverage relationship are robust to alternate specifications of the model.

The other robustness test involves separating out the index funds in the sample. Results for a sub sample of non-index funds may deviate from the full sample results as

[29] Ippolito (1992), Gruber (1996), Chevalier and Ellison (1997), Sirri and Tufano (1998)

media coverage may exert a stronger influence on non-index funds. To test whether inclusion of index funds has an impact on the empirical analysis presented earlier, I run the Heckman model regressions on sub samples of index and non-index funds. The results for the sub sample of index funds are presented in Tables 17 to 20 and display predictable characteristics such as loss of significance of past performance. It can be seen from these tables that only positive media coverage seems to have a significant impact on flows for index funds, and only the interactions with fund size is significant and not those with fund return. The results for regressions considering fund age are presented in Table 20 and show that the cumulative news article count is not significant for the index funds but the dummy variable for older funds is significant when interacted with the cumulative news variable. These results seem to be largely driven by Vanguard mutual funds which garner nearly 80% of the media coverage for the index funds. Moreover, over my sample period, the positive news stories for Vanguard outnumber the negative news stories by a factor of 10. This would contribute to the dominance of positive media coverage for large funds displayed by the index funds results. The corresponding results for the non-index sub sample are presented in Tables 21 to 24. These results are almost identical to the results for the full sample. This can be attributed to the fact that only 10% of the sample is comprised of index funds.

Further, to control for readership, circulation and other aspects specific to publications, I run the Heckman model regressions for two sub samples: the first sub sample is restricted to news articles appearing in *The Wall Street Journal* and the second sub sample includes news articles appearing in the other eleven publications. I find that the news variables are marginally stronger for the first sub sample than the second one.

Not surprisingly, the results suggest that the *Wall Street Journal* exerts a stronger influence on fund flows than the other publications. I also find that the mean fund size for the first sub sample is more than thrice the mean fund size for the second sub sample, indicating that the *Wall Street Journal* focuses more on larger funds.

The final robustness test involves the causality between mutual fund flows and media coverage. The empirical analysis presented earlier is based on causality from media coverage to fund flows. To confirm the correct direction of the causality, I employ vector auto regressions to conduct a test on the relation between media coverage and fund flows. Granger causality tests take into account the lagged values of the two series. The basic premise of a Granger causality test is to test whether variable *x* granger-causes variable *y*, that is, whether *x* leads *y* after controlling for past values of *y*. I conduct the Granger causality tests for both directions of the relation between fund flows and media coverage. I repeat the tests for three different lags of the two series. The results are presented in Table 25. Panel A presents results for the test of media coverage granger-causing fund flows. The Wald test statistics for causality show that there is significant causality in Model 1, that is, recent media coverage granger-causes fund flows. Panel B tests for fund flows granger-causing media coverage, and fails to find evidence for causality. These results provide support for recent media coverage leading to fund flows, and confirm the underlying assumption of causality in the empirical analysis presented in this dissertation.

Chapter 8

Conclusion

While there is a substantial body of literature on various aspects of mutual funds, there has been little research on media coverage of mutual funds. A number of studies have examined the effects of news on non-financial firms from both a theoretical and empirical perspective. However, there has not been a study that focuses on the attention and learning effects of media coverage for mutual funds. This dissertation contributes towards filling this gap in the mutual fund literature.

The principal focus of this dissertation is to investigate the pivotal role played by media coverage in the investment decisions of mutual fund investors. I study an underlying tenet of attention and learning models of mutual fund investors by examining sources of attention and learning, namely, news stories about funds. I propose that media coverage of mutual funds can influence financial decisions by capturing the attention of investors and facilitating in their learning about mutual funds. Specifically, I test whether the attention and learning effects of media coverage on investors are manifested through changes in investment in the fund. An advantage of examining attention and learning effects through the use of net fund flows is that one can obtain a more direct effect of news coverage than is possible with other types of assets, which have confounding valuation effects.

I find that media coverage of mutual funds has a significant effect on investor flows to the fund. Mutual funds with a news article in a month earn additional 1.2% net flows, on average. Further, it is evident that the media coverage affects investor learning in that media coverage does not have a uniform effect. Having positive media coverage in a month is associated with a positive 1.5% boost in fund flows, while having negative media coverage results in a negative 1% impact on fund flows. The significance of these numbers is further highlighted by the consideration that the average monthly fund flow for my sample is 1.08%. These results provide further evidence that media coverage can have significant economic effects on mutual funds. These results are also consistent with the hypothesis that investors learn about funds through media coverage and that this knowledge affects their investment behavior.

The database for media coverage of mutual funds comprising of nearly 10,000 news articles was constructed especially for this study. I examine the determinants of media coverage using several univariate and multivariate analyses. The results show that fund characteristics affect the likelihood of media coverage. Larger funds, those with extreme (high or low) performance, and those with more volatile returns are more likely to have news stories in general. Larger funds also have more positive news articles than smaller funds. I also find that fund size and past performance influence the impact of media coverage on mutual fund flows. Media coverage seems to play a more significant role for smaller and lesser known funds, and the fund's return performance enhances the impact of positive media coverage.

I find a differential in flows based on fund age in the effects of a news story, which is driven by the younger funds. I find that, as a fund ages and investors receive additional news about the fund, there are smaller effects from the news. Older funds do not receive as high fund flows as do younger funds, even controlling for the differences in sizes between the funds. I also construct a test to differentiate between the attention and learning effects of media coverage on mutual fund flows. I examine the flows into funds that were listed in a ranking article against funds with almost identical returns that were not listed in the ranking article. I find significant differences in flows between the two sets of funds, suggesting that just getting into a ranking article can provide significant attention effects for a fund.

I employ the Heckman model of self-selection to conduct the empirical analyses. Selection bias may occur when funds that receive media coverage have a greater propensity to receive higher inflows. I control for this potential endogeneity by using the Heckman selection model. I conduct an additional robustness check on the results using alternate specification of the regression model. I employ the Sirri and Tufano (1998) piecewise linear specification for return performance. I find that the economic and statistic significance of the news variables are robust for different model specifications.

Future research can include a comparative study of growth versus value funds. My dissertation performs an in-depth study of growth funds. An interesting parallel would be to examine investor reaction to value funds and compare the role of media coverage for value funds. There are many studies on *value* stocks earning higher returns than expensive *glamour* (growth) stocks. Lakonishok, et al. (1994) propose that the difference in the returns of value versus glamour stocks stems from investors' judgmental biases. It may be hence be interesting to examine the attention and learning effects of media coverage of value funds and the influence on investor decisions. Another avenue for future research would be to examine investor perception of the fund manager and the fund family in the context of media coverage of mutual funds.

Table 1. Fund Descriptive Statistics

The sample includes growth funds that existed over the 1994-2000 time period. This table provides descriptive statistics on the mutual funds with growth objective included in the sample. The 286 funds have a total of 406 share classes from the CRSP mutual fund database. The data for each of the share classes is combined by taking the value-weighted average across each share class connected to the fund. Total net assets is the total assets under management for the fund across all share classes. The flow is the total net flow to the fund over the month calculated as a percentage of the fund's assets. The return is the monthly total return per share for each fund. The expense ratio is the total operating expenses of the fund (including 12b-1 fees but excluding any load fees) calculated as a percentage of assets under management. The load fee is the total of all maximum front, deferred and rear-end load charges calculated as a percentage of assets under management. The fund age is calculated from the year the fund was organized. The mean, standard deviation, 10^{th} percentile, median and 90^{th} percentile are provided for each variable.
Panel A provides the fund descriptive statistics for the total sample. Panel B provides the same statistics for observations of funds with media coverage and observations of funds without media coverage. t-statistics are provided for the differences in the means of the fund characteristics for the sub samples with and without media coverage.

Variable	**Mean**	**Std Dev**	**10^{th} Pctl**	**Median**	**90^{th} Pctl**
Panel A. Total Sample					
Total Net Assets (in millions)	$1,692.00	$5,838.10	$22.12	$244.50	$3,320.50
Flow (monthly net percentage)	1.08%	18.40%	-2.71%	0.05%	4.34%
Total Return (monthly)	1.26%	4.80%	-4.14%	1.60%	6.42%
Expense Ratio	1.27%	1.71%	0.46%	1.08%	1.76%
Total Load	2.09%	2.47%	0.00%	3.91%	5.50%
Fund Age	15.52	16.13	3.00	9.00	41.00
Panel B. Sample divided by existence of media coverage					
Total Net Assets (in millions)					
Existence of media coverage	$4,460.76	$10,820.50	$21.80	$686.07	$11,427.01
Absence of media coverage	$1,030.65	$3,458.45	$22.16	$205.18	$1,985.25
t-statistic for the difference in means	21.29***				
Flow (monthly net percentage)					
Existence of media coverage	1.73%	13.16%	-2.66%	0.29%	5.79%
Absence of media coverage	0.92%	19.44%	-2.72%	0.00%	4.07%
t-statistic for the difference in means	3.39***				

Total Return (monthly)					
Existence of media coverage	1.19%	5.53%	-4.49%	1.57%	6.67%
Absence of media coverage	1.27%	4.60%	-4.08%	1.62%	6.38%
t-statistic for the difference in means	-0.91				
Expense Ratio (%)					
Existence of media coverage	1.53%	2.52%	0.40%	1.00%	2.08%
Absence of media coverage	1.20%	1.45%	0.48%	1.09%	1.72%
t-statistic for the difference in means	8.57***				
Total Load (%)					
Existence of media coverage	1.60%	2.47%	0.00%	0.00%	5.73%
Absence of media coverage	2.21%	2.45%	0.00%	0.46%	5.50%
t-statistic for the difference in means	-15.11***				
Fund Age (in years)					
Existence of media coverage	22.54	19.67	4.00	14.00	58.00
Absence of media coverage	13.85	14.68	3.00	8.25	36.00
t-statistic for the difference in means	28.21***				

*** significant at the 1% level ** significant at the 5% level * significant at the 10% level

Table 2. Circulation Data for Publications

This table provides circulation data for the 12 daily, weekly and monthly publications included in the study. The news data was constructed by searching for news stories in these 12 publications about the 286 funds in the sample, through the Dow Jones Retrieval Service (now called Factiva) for the 1994-2000 sample period. The circulations for newspapers are provided in Panel A (Source: *Audit Bureau of Circulations*). These include the highest circulation newspapers with the largest number of stories about mutual funds available from Factiva. The ranks in Panel A are based on circulations of all newspapers. The circulations for magazines are provided in Panel B (Source: *Media Distribution Services*). Each of these magazines had the largest number of stories about mutual funds in their category. The ranks in Panel B are based on circulations among magazines in the specified category.

Publication	Circulation (in '000)	Category	Rank
Panel A. Newspapers			
USA Today	2,666		1
Wall Street Journal	2,107		2
New York Times	1,681		3
Los Angeles Times	1,292		4
Washington Post	1,008		5
Boston Globe	708		12
Panel B. Magazines			
US News & World Report	2,019	News	3
Money	1,945	Personal Finance	1
Barron's	300	Personal Finance	6
Business Week	987	Business	2
Forbes	921	Business	3
Fortune	876	Business	4

Table 3. Frequency distribution of news articles

The sample includes growth funds that existed over the 1994-2000 time period. This table provides descriptive information on the 9,984 news articles portraying the 286 funds included in the sample. For the 12 daily, weekly and monthly publications, Panel A shows the name of the publication, the frequency of news mentions in that publication, and the percentage contribution of each publication to the total news stories. Panel B provides the frequency and percentage of articles by the posture of the article. Panel C shows frequency and percentage contribution of each type of news article. Appendix B describes the methodology of classification of the news articles.

	Frequency	%
Panel A. Publication		
Daily Publications:		
Boston Globe	1,023	10.25%
Los Angeles Times	203	2.03%
New York Times	61	0.61%
USA Today	1,427	14.29%
Washington Post	366	3.67%
Wall Street Journal	3,093	30.98%
Daily subtotal	6,173	61.83%
Weekly/Monthly Publications:		
Barron's	1,383	13.85%
Business Week	776	7.77%
Forbes	319	3.20%
Fortune	251	2.51%
Money	733	7.34%
US News & World Report	349	3.50%
Weekly/Monthly subtotal	3,811	38.17%
Total News Articles	9,984	100.00%
Panel B. Posture of news story		
Positive	2,774	27.78%
Neutral-Positive	898	8.99%

Neutral	5,024	50.32%
Neutral-Negative	307	3.08%
Negative	981	9.83%
Total News Articles	9,984	100.00%
Panel C. Type of news story		
Feature	460	4.61%
Interview	181	1.81%
Ranking	297	2.97%
Performance	472	4.73%
Mention	3811	38.17%
Tables	4763	47.71%
Total News Articles	9,984	100.00%

Table 4. Descriptive statistics of news articles

The sample includes growth funds that existed over the 1994-2000 time period. This table shows descriptive statistics for the news data for the sample funds. Panel A provides the descriptive information for the existence of a news article for a fund in a month. The existence of a particular category of media coverage is denoted by the relevant dummy variable. The news dummy variable takes on value of one if there existed at least one news article in that category in a given month; else the dummy takes on value of zero. Panel B shows the descriptive statistics for the actual number of news articles in each category of classification for a fund in a month.

	Mean	Std Dev	Maximum	Median
Panel A. Existence of news article in a month				
Dummy for whether fund had:				
Any news article	0.19	0.40	1	0
Positive news article	0.10	0.30	1	0
Neutral news article	0.11	0.31	1	0
Negative news article	0.04	0.19	1	0
Panel B. Frequency of news article in a month				
Number of articles	0.42	1.31	24	0
Number of positive articles	0.15	0.60	12	0
Number of neutral articles	0.21	0.90	22	0
Number of negative articles	0.05	0.32	9	0

Table 5. Correlation Matrix

The sample includes growth funds that existed over the 1994-2000 time period. This table shows the correlation between the variables used in the various analyses. Detailed definitions of all fund and news variables are provided in Chapter 4.

Variable	Market Flow	Monthly Fund Flow	Total Net Assets	Monthly Fund Return	Annual Fund Return	Return Volatility	Expense Ratio	Load Fees
Market Flow	1.000	0.093	-0.091	0.020	-0.100	-0.407	-0.023	-0.021
Monthly Fund Flow	0.093	1.000	-0.010	0.021	0.053	-0.031	-0.008	-0.005
Total Net Assets	-0.091	-0.010	1.000	0.004	0.036	0.030	-0.099	0.025
Monthly Fund Return	0.020	0.021	0.004	1.000	0.306	-0.050	-0.044	0.000
Annual Fund Return	-0.100	0.053	0.036	0.306	1.000	0.076	-0.071	0.010
Return Volatility	-0.407	-0.031	0.030	-0.050	0.076	1.000	0.275	0.057
Expense Ratio	-0.023	-0.008	-0.099	-0.044	-0.071	0.275	1.000	0.073
Load Fees	-0.021	-0.005	0.025	0.000	0.010	0.057	0.073	1.000
Fund Age	-0.073	-0.052	0.214	-0.015	-0.032	0.062	0.103	0.065
# news articles	0.002	0.016	0.418	-0.001	0.054	0.079	0.049	-0.102
Cumulative # news articles	-0.159	-0.016	0.594	-0.018	0.014	0.225	0.131	-0.116
# positive news articles	0.019	0.028	0.231	0.012	0.148	0.048	-0.012	-0.104
# neutral news articles	-0.009	0.010	0.436	0.002	0.030	0.030	-0.014	-0.075
# negative news articles	-0.001	-0.016	0.048	-0.030	-0.138	0.149	0.257	-0.011
Dummy for existence of any news	0.016	0.018	0.232	-0.007	0.031	0.065	0.076	-0.096
Dummy for existence of positive news	0.011	0.029	0.192	0.007	0.124	0.039	-0.006	-0.102
Dummy for existence of neutral news	0.006	0.011	0.292	0.002	0.028	0.035	0.009	-0.072
Dummy for existence of negative news	0.002	-0.018	0.043	-0.040	-0.125	0.119	0.193	-0.028
Performance news articles	0.008	0.005	0.220	0.003	0.037	0.015	0.008	-0.060
Ranking news articles	-0.003	0.011	0.120	0.008	0.114	0.067	-0.004	-0.047
Mentions news articles	-0.018	0.013	0.406	0.002	0.046	0.038	-0.018	-0.107
Tables news articles	0.019	0.012	0.278	-0.004	0.021	0.084	0.096	-0.051
Feature news articles	-0.005	0.004	0.152	-0.013	0.027	0.037	0.036	-0.054
Interview news articles	0.001	0.010	0.040	0.004	0.000	0.017	-0.007	-0.020

Table 5. Correlation Matrix – continued.

The sample includes growth funds that existed over the 1994-2000 time period. This table shows the correlations between the variables of media coverage and fund characteristics used in the various analyses. Detailed definitions of all fund and news variables are provided in Chapter 4.

Variable	Fund Age	# news articles	Cumulative # news articles	# positive news articles	# neutral news articles	# negative news articles	Dummy for existence of any news	Dummy for existence of positive news
Market Flow	-0.073	0.002	-0.159	0.019	-0.009	-0.001	0.016	0.011
Monthly Fund Flow	-0.052	0.016	-0.016	0.028	0.010	-0.016	0.018	0.029
Total Net Assets	0.214	0.418	0.594	0.231	0.436	0.048	0.232	0.192
Monthly Fund Return	-0.015	-0.001	-0.018	0.012	0.002	-0.030	-0.007	0.007
Annual Fund Return	-0.032	0.054	0.014	0.148	0.030	-0.138	0.031	0.124
Return Volatility	0.062	0.079	0.225	0.048	0.030	0.149	0.065	0.039
Expense Ratio	0.103	0.049	0.131	-0.012	-0.014	0.257	0.076	-0.006
Load Fees	0.065	-0.102	-0.116	-0.104	-0.075	-0.011	-0.096	-0.102
Fund Age	1.000	0.162	0.239	0.130	0.115	0.096	0.212	0.147
# news articles	0.162	1.000	0.612	0.703	0.861	0.344	0.653	0.548
Cumulative # news articles	0.239	0.612	1.000	0.337	0.582	0.229	0.398	0.288
# positive news articles	0.130	0.703	0.337	1.000	0.329	0.080	0.525	0.786
# neutral news articles	0.115	0.861	0.582	0.329	1.000	0.087	0.477	0.245
# negative news articles	0.096	0.344	0.229	0.080	0.087	1.000	0.339	0.078
Dummy for existence of any news	0.212	0.653	0.398	0.525	0.477	0.339	1.000	0.668
Dummy for existence of positive news	0.147	0.548	0.288	0.786	0.245	0.078	0.668	1.000
Dummy for existence of neutral news	0.170	0.619	0.423	0.278	0.679	0.095	0.702	0.249
Dummy for existence of negative news	0.106	0.307	0.216	0.076	0.097	0.828	0.410	0.076
Performance news articles	0.045	0.440	0.240	0.288	0.376	0.201	0.218	0.235
Ranking news articles	0.017	0.276	0.132	0.411	0.111	0.050	0.200	0.292
Mentions news articles	0.140	0.801	0.525	0.589	0.692	0.222	0.506	0.440
Tables news articles	0.132	0.822	0.488	0.506	0.740	0.325	0.564	0.412
Feature news articles	0.072	0.396	0.235	0.348	0.286	0.160	0.235	0.266
Interview news articles	0.048	0.176	0.098	0.130	0.158	0.031	0.168	0.114

Table 5. Correlation Matrix – continued.

The sample includes growth funds that existed over the 1994-2000 time period. This table shows the correlation between the variables used in the various analyses. Detailed definitions of all fund and news variables are provided in Chapter 4.

Variable	Dummy for existence of neutral news	Dummy for existence of negative news	Perfor-mance news articles	Ranking news articles	Mentions news articles	Tables news articles	Feature news articles	Interview news articles
Market Flow	0.006	0.002	0.008	-0.003	-0.018	0.019	-0.005	0.001
Monthly Fund Flow	0.011	-0.018	0.005	0.011	0.013	0.012	0.004	0.010
Total Net Assets	0.292	0.043	0.220	0.120	0.406	0.278	0.152	0.040
Monthly Fund Return	0.002	-0.040	0.003	0.008	0.002	-0.004	-0.013	0.004
Annual Fund Return	0.028	-0.125	0.037	0.114	0.046	0.021	0.027	0.000
Return Volatility	0.035	0.119	0.015	0.067	0.038	0.084	0.037	0.017
Expense Ratio	0.009	0.193	0.008	-0.004	-0.018	0.096	0.036	-0.007
Load Fees	-0.072	-0.028	-0.060	-0.047	-0.107	-0.051	-0.054	-0.020
Fund Age	0.170	0.106	0.045	0.017	0.140	0.132	0.072	0.048
# news articles	0.619	0.307	0.440	0.276	0.801	0.822	0.396	0.176
Cumulative # news articles	0.423	0.216	0.240	0.132	0.525	0.488	0.235	0.098
# positive news articles	0.278	0.076	0.288	0.411	0.589	0.506	0.348	0.130
# neutral news articles	0.679	0.097	0.376	0.111	0.692	0.740	0.286	0.158
# negative news articles	0.095	0.828	0.201	0.050	0.222	0.325	0.160	0.031
Dummy for existence of any news	0.702	0.410	0.218	0.200	0.506	0.564	0.235	0.168
Dummy for existence of positive news	0.249	0.076	0.235	0.292	0.440	0.412	0.266	0.114
Dummy for existence of neutral news	1.000	0.112	0.139	0.109	0.552	0.501	0.204	0.202
Dummy for existence of negative news	0.112	1.000	0.164	0.040	0.211	0.282	0.138	0.042
Performance news articles	0.139	0.164	1.000	0.124	0.255	0.255	0.164	-0.004
Ranking news articles	0.109	0.040	0.124	1.000	0.160	0.133	0.054	0.038
Mentions news articles	0.552	0.211	0.255	0.160	1.000	0.391	0.262	0.116
Tables news articles	0.501	0.282	0.255	0.133	0.391	1.000	0.190	0.063
Feature news articles	0.204	0.138	0.164	0.054	0.262	0.190	1.000	0.080
Interview news articles	0.202	0.042	-0.004	0.038	0.116	0.063	0.080	1.000

Table 6. Media Coverage and Fund Characteristics - Univariate Analysis

The sample includes growth funds that existed over the 1994-2000 time period. This table provides the results of univariate analyses of media coverage for fund characteristics. For each panel, the funds in the sample are divided into quintiles after sorting based on the relevant fund characteristic. For each quintile, the average number of news articles per month is calculated. The first column provides the average number of news articles for overall media coverage. The other three columns provide the averages for positive, neutral or negative news articles based on the posture of the news article towards the fund portrayed. Panel A focuses on Size, measured as Total Net Assets. Panel B focuses on Performance, measured as Return previous year. Panel C focuses on the volatility of the fund return for the previous year.

	Any Articles	Positive Articles	Neutral Articles	Negative Articles
Panel A. Average number of news articles by fund size				
Size Quintile				
Largest	1.12	0.41	0.64	0.06
4	0.31	0.13	0.15	0.03
3	0.20	0.09	0.08	0.03
2	0.22	0.07	0.12	0.03
Smallest	0.26	0.07	0.07	0.12
Panel B. Average number of news articles by fund performance				
Performance Quintile				
Highest	0.59	0.26	0.29	0.04
4	0.38	0.16	0.20	0.02
3	0.37	0.15	0.19	0.03
2	0.37	0.14	0.18	0.05
Lowest	0.39	0.07	0.20	0.13
Panel C. Average number of news articles by return volatility				
Volatility Quintile				
Highest	0.60	0.20	0.27	0.12
4	0.35	0.12	0.18	0.05
3	0.43	0.14	0.25	0.04
2	0.35	0.15	0.18	0.03
Lowest	0.38	0.16	0.18	0.04

Table 7. Fund Performance and Posture of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table showcases a study of the influence of fund performance on the posture of media coverage. The study highlights the differences in proportions of news stories with positive/negative bias following positive/negative fund performance. The sample is divided into positive and negative bins based on whether the past fund returns were positive or negative. Then for each bin, the proportions of positive and negative news stories are obtained. This process is repeated for two different horizons of past return: previous month and previous year. Panel A provides results of this analysis for the total sample. Panel B considers the smaller half of the sample, obtained by restricting the fund total net assets (TNA) to below the sample median. Thus the results in Panel B will help isolate the sensitivity of the posture of media coverage to past performance of smaller funds.

Panel A: Total Sample

# total news stories	# positive news stories	# negative news stories
9984	3672	1288
	Proportion of Positive News	**Proportion of Negative News**
Return Previous Month:		
Positive Bin	0.40	0.10
Negative Bin	0.32	0.17
Return Previous Year:		
Positive Bin	0.41	0.09
Negative Bin	0.11	0.40

Panel B: Smaller Half of Sample

# total news stories	# positive news stories	# negative news stories
2806	937	798
	Proportion of Positive News	**Proportion of Negative News**
Return Previous Month:		
Positive Bin	0.40	0.22
Negative Bin	0.23	0.38
Return Previous Year:		
Positive Bin	0.42	0.20
Negative Bin	0.08	0.57

Table 8. Media Coverage and Fund Characteristics – Multivariate Analysis

The sample includes growth funds that existed over the 1994-2000 time period. This table provides the results of regressions using binary probit model in which the response variable is the flag for the existence of media coverage in a month, with two response levels of 1 and 0. The probability modeled is the existence of media coverage denoted by the response level of 1. Panel A presents the results from the regression for modeling the probability of any media coverage. Panel B presents the results from the regression run separately for modeling the probability of positive media coverage. Panel C presents the results from the regression run separately for modeling the probability of negative media coverage. The independent variables include the natural logarithm of the previous month's total net assets under management (TNA), the return on the fund for the previous month, the return on the fund for the previous year, the volatility of the fund return for the previous year, and the interaction term between the return for previous year and a dummy variable for small funds. The *small* dummy variable takes on a value of one if the fund's TNA is smaller than the TNA of the median fund in the sample, and takes on a value of zero if the fund's TNA exceeds the median fund's TNA. p-values are provided in parentheses below the coefficient estimates.

	Model 1 **Any news**	**Model 2** **Positive news**	**Model 3** **Negative news**
Intercept	-1.796	-2.703	-1.628
	(0.000)	(0.000)	(0.000)
Log TNA previous month	0.127	0.183	-0.060
	(0.000)	(0.000)	(0.000)
Return previous month	0.549	0.304	0.324
	(0.008)	(0.233)	(0.282)
Return previous year	0.173	0.834	-0.960
	(0.018)	(0.000)	(0.000)
Return volatility	4.729	2.443	8.872
	(0.000)	(0.000)	(0.000)
Small dummy * Return previous year	-0.4425	0.488	-1.003
	(0.000)	(0.000)	(0.000)
Likelihood Ratio	1,052.33	1,131.54	628.40
	(0.000)	(0.000)	(0.000)
Score	1,095.14	1,142.92	919.38
	(0.000)	(0.000)	(0.000)

Wald	962.85	951.12	572.78
	(0.000)	(0.000)	(0.000)

Table 9. Relation of Absolute Fund Flows to Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the absolute value of percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variable included in Model 1 is the actual number of news stories portraying the fund each month. The media coverage variable included in Model 2 is the actual number of news stories portraying the fund each month, weighted by the proportion of the circulation of each periodical to the total circulation of all the twelve periodicals. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.020	-0.016
	(0.908)	(0.927)
Market Flow	0.083	0.073
	(0.275)	(0.342)
Flow (t-1)	0.116	0.116
	(0.000)	(0.000)
Return previous year	0.024	0.024
	(0.000)	(0.000)
Log TNA (t-1)	-0.003	-0.003
	(0.664)	(0.656)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.313	-0.304
	(0.000)	(0.000)
Volatility of return previous year	-0.215	-0.200
	(0.289)	(0.324)
Number of news articles	0.002	
	(0.000)	

Circulation-weighted number of news articles		0.029
		(0.000)
Sigma	0.087	0.086
	(0.000)	(0.000)
Rho	0.388	0.372
	(0.513)	(0.539)
Log likelihood	-5921	-5919
N	23466	23466

Table 10. Relation of Directional Fund Flows to Posture of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year. The media coverage variables included in Model 1 are the dummy variables for the existence of news stories in each period, classified by the posture of the article towards the fund portrayed. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed. p-values are provided in parentheses below the coefficient estimates.

	Model 1	Model 2
Intercept	-0.210	-0.213
	(0.290)	(0.286)
Market Flow	0.418	0.403
	(0.000)	(0.000)
Flow (t-1)	0.006	0.007
	(0.694)	(0.626)
Return previous year	0.080	0.079
	(0.000)	(0.000)
Log TNA (t-1)	0.005	0.005
	(0.513)	(0.497)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.099	-0.096
	(0.151)	(0.168)
Volatility of return previous year	0.226	0.225
	(0.340)	(0.345)
Dummy for whether the fund has:		
Positive media coverage	0.015	
	(0.000)	

Neutral media coverage	0.007	
	(0.051)	
Negative media coverage	-0.010	
	(0.009)	
Number of positive news articles		0.006
		(0.000)
Number of neutral news articles		0.002
		(0.043)
Number of negative news articles		-0.006
		(0.005)
Sigma	0.119	0.121
	(0.015)	(0.015)
Rho	0.689	0.702
	(0.028)	(0.019)
Log likelihood	-6260	-6262
N	23466	23466

Table 11. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Size

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.

The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the log of the fund's TNA for the previous period and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the log of the fund's TNA for the previous period. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.185	-0.215
	(0.337)	(0.272)
Market Flow	0.418	0.373
	(0.000)	(0.000)
Flow (t-1)	0.007	-0.001
	(0.613)	(0.927)
Return previous year	0.091	0.075
	(0.000)	(0.000)
Log TNA (t-1)	0.006	0.008
	(0.389)	(0.302)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.190	-0.100
	(0.007)	(0.167)
Volatility of return previous year	0.121	0.167
	(0.596)	(0.473)
Number of news articles	0.014	
	(0.000)	

Log TNA (t-1) * Number of news articles	-0.001	
	(0.000)	
Number of positive news articles		0.031
		(0.000)
Number of neutral news articles		0.013
		(0.000)
Number of negative news articles		0.003
		(0.457)
Log TNA (t-1) * Number of positive news articles		-0.003
		(0.000)
Log TNA (t-1) * Number of neutral news articles		-0.001
		(0.000)
Log TNA (t-1) * Number of negative news articles		-0.001
		(0.112)
Sigma	0.109	0.115
	(0.009)	(0.013)
Rho	0.611	0.665
	(0.121)	(0.049)
Log likelihood	-6260	-6237
N	23466	23466

Table 12. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Performance

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.

The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the return to the fund for the previous year and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the return to the fund for the previous year. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.218	-0.223
	(0.276)	(0.264)
Market Flow	0.441	0.440
	(0.000)	(0.000)
Flow (t-1)	0.012	0.001
	(0.407)	(0.979)
Return previous year	0.105	0.105
	(0.000)	(0.000)
Log TNA (t-1)	0.005	0.006
	(0.473)	(0.457)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.156	-0.121
	(0.025)	(0.083)
Volatility of return previous year	0.212	0.214
	(0.374)	(0.371)
Number of news articles	0.004	
	(0.000)	

Return previous year * Number of news articles	-0.005	
	(0.022)	
Number of positive news articles		0.002
		(0.221)
Number of neutral news articles		0.006
		(0.000)
Number of negative news articles		-0.004
		(0.055)
Return previous year * Number of positive news articles		0.013
		(0.002)
Return previous year * Number of neutral news articles		-0.020
		(0.000)
Return previous year * Number of negative news articles		-0.034
		(0.000)
Sigma	0.120	0.121
	(0.015)	(0.016)
Rho	0.697	0.708
	(0.022)	(0.015)
Log likelihood	-6272	-6232
N	23466	23466

Table 13. Relation of Directional Fund Flows to Media Coverage considering Age of Fund

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year. The media coverage variables included in Model 1 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age variable and the number of news articles and between the age variable and the cumulative number of news articles. The media coverage variables included in Model 2 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age group dummy variables and the number of news articles and between the age group dummy variables and the cumulative number of news articles. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.200	-0.156
	(0.315)	(0.399)
Market Flow	0.270	0.267
	(0.004)	(0.005)
Flow (t-1)	0.005	-0.008
	(0.706)	(0.593)
Return previous year	0.088	0.083
	(0.000)	(0.000)
Log TNA (t-1)	0.006	0.004
	(0.459)	(0.542)
Age	-0.001	-0.001
	(0.017)	(0.062)
Expense ratio (t-1)	-0.145	-0.148
	(0.037)	(0.032)
Volatility of return previous year	0.255	0.110
	(0.283)	(0.615)

Number of news articles	0.005	0.002
	(0.000)	(0.127)
Cumulative number of news articles	-0.035	-0.022
	(0.000)	(0.005)
Age * Number of news articles	-0.001	
	(0.002)	
Age * Cumulative number of news articles	0.001	
	(0.012)	
Young funds * Number of news articles		0.012
		(0.000)
Young funds * Cumulative number of news articles		-0.011
		(0.205)
Old funds * Number of news articles		-0.001
		(0.336)
Old funds * Cumulative number of news articles		0.005
		(0.495)
Sigma	0.119	0.107
	(0.015)	(0.006)
Rho	0.690	0.598
	(0.027)	(0.128)
Log likelihood	-6259	-6195
N	23466	23466

Table 14. The Impact of Ranking Articles on Fund Flows

The sample includes growth funds that existed over the 1994-2000 time period. This table showcases the differences in flows between two sets of matched funds with similar returns, only one of which has been portrayed in a Ranking article. Each fund included in a ranking article is paired with the fund closest to it in terms of return, the difference in the two returns being no more than 0.005. The table shows the summary statistics for each set of the matched funds. The t-test statistics are provided for the differences in flows and differences in returns between the two funds in each pair.

	Flow of Ranked Fund	**Flow of Non-Ranked Fund**	**Return of Ranked Fund**	**Return of Non-Ranked Fund**
Mean	3.41 %	0.27 %	15.15 %	15.01 %
Standard Deviation	9.79 %	5.04 %	17.78 %	17.77 %
t-statistic for the difference	3.61***		0.07	

*** significant at the 1% level ** significant at the 5% level * significant at the 10% level

Table 15. Relation of Fund Flows to Media Coverage employing the Piecewise Linear Specification

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from alternate specification of the regression model to check for robustness of the regression results. The piecewise linear specification for the fund flow-performance relation proposed by Sirri and Tufano (1998) is employed. The independent variables include log of the fund's total net assets under management (TNA) for the previous period, the aggregate flows for all funds in the same investment category, the volatility of the previous year's monthly returns, the level of total fees (expense ratio plus amortized load) charged by the fund for an investor with a seven-year holding period, and measures of fractional performance rank of the fund in the previous year.

A fund's fractional RANK represents its percentile performance relative to other funds with the same investment objective in the same period, and ranges from 0 to 1. This table presents the results for the model with four kinks, where the coefficients on fractional ranks are estimated using a piecewise linear regression framework over five quintiles. The coefficients on these piecewise linear decompositions of fractional ranks represent the slope of the performance-flow relationship over their range of sensitivity. These regressions are run on cross-sectionally for each of the 84 months of the sample period, and the Newey-West (1987) t-statistics are calculated for the coefficients from the vector of monthly results, employing the Fama and Macbeth (1973) technique. The t-statistics are given in parentheses below the coefficient estimates. The average adjusted R-squared values are also provided.

The media coverage variable included in Model 1 is the dummy variable for the existence of any news article portraying the fund in each period. The media coverage variable included in Model 2 is the total number of news stories in each period potraying the fund.

	Model 1	**Model 2**
Intercept	0.005	0.007
	(0.64)	(0.89)
Log Lag TNA	-0.004	-0.004
	(-5.05)***	(-5.16)***
Flows to fund category	0.999	1.005
	(3.87)***	(3.71)***
Std. Dev. of returns	-0.367	-0.372
	(-3.74)***	(-3.70)***
Total fees	0.003	0.003
	(1.36)	(1.44)
Breakdown of RANK:		
Bottom performance quintile	0.045	0.043

	(2.54)**	(2.45)**
4th performance quintile	0.046	0.043
	(3.66)***	(3.92)***
3rd performance quintile	0.034	0.040
	(2.41)**	(3.20)***
2nd performance quintile	0.022	0.016
	(1.60)	(1.17)
Top performance quintile	0.125	0.128
	(3.60)***	(3.74)***
Dummy for whether the fund has:	0.012	
Any media coverage	(6.43)***	
Number of news articles		0.004
		(5.81)***
Adj. R^2 (%)	7.01	7.22

*** significant at the 1% level ** significant at the 5% level * significant at the 10% level

Table 16. Relation of Fund Flows to Media Coverage employing the Piecewise Linear Specification considering Posture of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from alternate specification of the regression model to check for robustness of the regression results. The piecewise linear specification for the fund flow-performance relation proposed by Sirri and Tufano (1998) is employed. The independent variables include log of the fund's total net assets under management (TNA) for the previous period, the aggregate flows for all funds in the same investment category, the volatility of the previous year's monthly returns, the level of total fees (expense ratio plus amortized load) charged by the fund for an investor with a seven-year holding period, and measures of fractional performance rank of the fund in the previous year. A fund's fractional RANK represents its percentile performance relative to other funds with the same investment objective in the same period, and ranges from 0 to 1. This table presents the results for the model with four kinks, where the coefficients on fractional ranks are estimated using a piecewise linear regression framework over five quintiles. The coefficients on these piecewise linear decompositions of fractional ranks represent the slope of the performance-flow relationship over their range of sensitivity. These regressions are run on cross-sectionally for each of the 84 months of the sample period, and the Newey-West (1987) t-statistics are calculated for the coefficients from the vector of monthly results, employing the Fama and Macbeth (1973) technique. The t-statistics are given in parentheses below the coefficient estimates. The average adjusted R-squared values are also provided. The media coverage variables included in Model 1 are the dummy variables for the existence of news stories in each period, classified by the posture of the article towards the fund portrayed. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed.

	Model 1	**Model 2**
Intercept	0.011	0.011
	(1.47)	(1.50)
Log Lag TNA	-0.004	-0.004
	(-5.31)***	(-5.28)***
Flows to fund category	0.994	1.004
	(3.73)***	(3.66)***
Std. Dev. of returns	-0.358	-0.350
	(-3.75)***	(-3.72)***
Total fees	0.004	0.004
	(1.97)**	(1.84)*
Breakdown of RANK:		
Bottom performance quintile	0.028	0.025
	(1.55)	(1.40)

4th performance quintile	0.036	0.039
	(3.18)***	(3.66)***
3rd performance quintile	0.039	0.039
	(3.09)***	(3.29)***
2nd performance quintile	0.020	0.018
	(1.53)	(1.37)
Top performance quintile	0.122	0.121
	(3.56)***	(3.57)***
Dummy for whether the fund has:		
Positive media coverage	0.018	
	(4.35)***	
Neutral media coverage	0.012	
	(4.20)***	
Negative media coverage	-0.013	
	(-3.61)***	
Number of positive news articles		0.009
		(5.58)***
Number of neutral news articles		0.005
		(2.92)***
Number of negative news articles		-0.011
		(-3.33)***
Adj. R^2 (%)	7.61	7.88

*** significant at the 1% level ** significant at the 5% level * significant at the 10% level

Table 17. Relation of Directional Fund Flows to Posture of Media Coverage for Index Funds

The sample includes index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.

The media coverage variables included in Model 1 are the dummy variables for the existence of news stories in each period, classified by the posture of the article towards the fund portrayed. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	0.227	0.072
	(0.806)	(0.932)
Market Flow	0.361	0.355
	(0.013)	(0.015)
Flow (t-1)	0.386	0.387
	(0.000)	(0.000)
Return previous year	0.042	0.028
	(0.656)	(0.752)
Log TNA (t-1)	-0.017	-0.006
	(0.788)	(0.915)
Age	0.001	0.001
	(0.507)	(0.684)
Expense ratio (t-1)	-1.691	-1.357
	(0.272)	(0.381)
Volatility of return previous year	0.552	0.010
	(0.866)	(0.997)
Dummy for whether the fund has:		
Positive media coverage	0.009	

	(0.121)	
Neutral media coverage	-0.004	
	(0.517)	
Negative media coverage	-0.006	
	(0.358)	
Number of positive news articles		0.005
		(0.031)
Number of neutral news articles		-0.001
		(0.542)
Number of negative news articles		-0.004
		(0.309)
Sigma	0.066	0.040
	(0.760)	(0.506)
Rho	-0.811	-0.247
	(0.557)	(0.966)
Log likelihood	-132.69	-133.73
N	2414	2414

Table 18. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Size for Index Funds

The sample includes index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the log of the fund's TNA for the previous period and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the log of the fund's TNA for the previous period. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	0.141	0.388
	(0.871)	(0.731)
Market Flow	0.367	0.305
	(0.013)	(0.036)
Flow (t-1)	0.407	0.379
	(0.000)	(0.000)
Return previous year	0.036	0.056
	(0.687)	(0.633)
Log TNA (t-1)	-0.011	-0.027
	(0.861)	(0.726)
Age	0.001	0.001
	(0.876)	(0.833)
Expense ratio (t-1)	-1.257	-1.256
	(0.407)	(0.421)
Volatility of return previous year	0.267	1.049
	(0.931)	(0.794)
Number of news articles	0.001	
	(0.926)	

Log TNA (t-1) * Number of news articles	0.001	
	(0.960)	
Number of positive news articles		0.026
		(0.016)
Number of neutral news articles		-0.008
		(0.289)
Number of negative news articles		-0.018
		(0.227)
Log TNA (t-1) * Number of positive news articles		-0.002
		(0.032)
Log TNA (t-1) * Number of neutral news articles		0.001
		(0.265)
Log TNA (t-1) * Number of negative news articles		0.002
		(0.297)
Sigma	0.050	0.108
	(0.751)	(0.723)
Rho	-0.625	-0.935
	(0.839)	(0.013)
Log likelihood	-136.85	-129.29
N	2414	2414

Table 19. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Performance for Index Funds

The sample includes index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the return to the fund for the previous year and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the return to the fund for the previous year. p-values are provided in parentheses below the coefficient estimates.

	Model 1	Model 2
Intercept	0.150	0.047
	(0.864)	(0.955)
Market Flow	0.368	0.353
	(0.012)	(0.016)
Flow (t-1)	0.407	0.385
	(0.000)	(0.000)
Return previous year	0.038	0.029
	(0.678)	(0.743)
Log TNA (t-1)	-0.011	-0.005
	(0.854)	(0.937)
Age	0.001	0.001
	(0.878)	(0.653)
Expense ratio (t-1)	-1.258	-1.467
	(0.406)	(0.355)
Volatility of return previous year	0.296	-0.084
	(0.925)	(0.978)
Number of news articles	0.001	
	(0.769)	

Return previous year * Number of news articles	0.001	
	(0.926)	
Number of positive news articles		0.008
		(0.137)
Number of neutral news articles		-0.001
		(0.613)
Number of negative news articles		-0.005
		(0.606)
Return previous year * Number of positive news articles		-0.014
		(0.525)
Return previous year * Number of neutral news articles		0.002
		(0.731)
Return previous year * Number of negative news articles		0.002
		(0.972)
Sigma	0.051	0.039
	(0.756)	(0.014)
Rho	-0.655	-0.065
	(0.816)	(0.992)
Log likelihood	-136.84	-133.52
N	2414	2414

Table 20. Relation of Directional Fund Flows to Media Coverage considering Age of Fund for Index Funds

The sample includes index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.

The media coverage variables included in Model 1 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age variable and the number of news articles and between the age variable and the cumulative number of news articles. The media coverage variables included in Model 2 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age group dummy variables and the number of news articles and between the age group dummy variables and the cumulative number of news articles. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.020	0.249
	(0.981)	(0.793)
Market Flow	0.400	0.341
	(0.015)	(0.038)
Flow (t-1)	0.398	0.379
	(0.000)	(0.000)
Return previous year	0.030	0.042
	(0.735)	(0.667)
Log TNA (t-1)	-0.002	-0.021
	(0.970)	(0.754)
Age	0.001	0.001
	(0.556)	(0.293)
Expense ratio (t-1)	-0.927	-3.013
	(0.550)	(0.064)
Volatility of return previous year	-0.092	0.616
	(0.975)	(0.854)

Number of news articles	0.008	0.019
	(0.085)	(0.001)
Cumulative number of news articles	-0.025	-0.008
	(0.143)	(0.557)
Age * Number of news articles	0.001	
	(0.554)	
Age * Cumulative number of news articles	0.002	
	(0.019)	
Young funds * Number of news articles		-0.015
		(0.015)
Young funds * Cumulative number of news articles		0.006
		(0.707)
Old funds * Number of news articles		-0.019
		(0.001)
Old funds * Cumulative number of news articles		0.023
		(0.091)
Sigma	0.039	0.069
	(0.005)	(0.762)
Rho	0.057	-0.833
	(0.993)	(0.493)
Log likelihood	-133.45	-128.65
N	2414	2414

Table 21. Relation of Directional Fund Flows to Posture of Media Coverage for Non-Index Funds

The sample includes non-index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.

The media coverage variables included in Model 1 are the dummy variables for the existence of news stories in each period, classified by the posture of the article towards the fund portrayed. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.198	-0.192
	(0.294)	(0.308)
Market Flow	0.425	0.407
	(0.000)	(0.000)
Flow (t-1)	0.002	0.003
	(0.891)	(0.833)
Return previous year	0.081	0.081
	(0.000)	(0.000)
Log TNA (t-1)	0.004	0.004
	(0.550)	(0.564)
Age	0.001	0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.103	-0.102
	(0.154)	(0.162)
Volatility of return previous year	0.279	0.260
	(0.036)	(0.341)
Dummy for whether the fund has:		
Positive media coverage	0.015	

	(0.000)	
Neutral media coverage	0.007	
	(0.052)	
Negative media coverage	-0.011	
	(0.010)	
Number of positive news articles		0.006
		(0.000)
Number of neutral news articles		0.002
		(0.014)
Number of negative news articles		-0.006
		(0.012)
Sigma	0.118	0.118
	(0.007)	(0.007)
Rho	0.654	0.653
	(0.044)	(0.045)
Log likelihood	-5800	-5800
N	21052	21052

Table 22. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Size for Non-Index Funds

The sample includes non-index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the log of the fund's TNA for the previous period and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the log of the fund's TNA for the previous period. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.174	-0.202
	(0.340)	(0.275)
Market Flow	0.426	0.382
	(0.000)	(0.000)
Flow (t-1)	0.003	-0.005
	(0.827)	(0.727)
Return previous year	0.093	0.077
	(0.000)	(0.000)
Log TNA (t-1)	0.006	0.007
	(0.382)	(0.293)
Age	0.001	0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.196	-0.104
	(0.007)	(0.166)
Volatility of return previous year	0.155	0.213
	(0.560)	(0.427)
Number of news articles	0.015	
	(0.000)	

Log TNA (t-1) * Number of news articles	-0.001	
	(0.000)	
Number of positive news articles		0.031
		(0.000)
Number of neutral news articles		0.015
		(0.000)
Number of negative news articles		0.003
		(0.361)
Log TNA (t-1) * Number of positive news articles		-0.003
		(0.000)
Log TNA (t-1) * Number of neutral news articles		-0.001
		(0.000)
Log TNA (t-1) * Number of negative news articles		-0.001
		(0.086)
Sigma	0.109	0.113
	(0.003)	(0.006)
Rho	0.570	0.622
	(0.158)	(0.080)
Log likelihood	-5798	-5777
N	21052	21052

Table 23. Relation of Directional Fund Flows to Interaction of Media Coverage and Fund Performance for Non-Index Funds

The sample includes non-index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the return to the fund for the previous year and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the return to the fund for the previous year. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.200	-0.212
	(0.291)	(0.264)
Market Flow	0.443	0.451
	(0.000)	(0.000)
Flow (t-1)	0.008	-0.003
	(0.604)	(0.827)
Return previous year	0.107	0.108
	(0.000)	(0.000)
Log TNA (t-1)	0.005	0.005
	(0.523)	(0.478)
Age	0.001	0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.162	-0.122
	(0.026)	(0.091)
Volatility of return previous year	0.243	0.272
	(0.376)	(0.323)
Number of news articles	0.004	
	(0.000)	

Return previous year * Number of news articles	-0.006	
	(0.030)	
Number of positive news articles		0.002
		(0.256)
Number of neutral news articles		0.007
		(0.000)
Number of negative news articles		-0.004
		(0.064)
Return previous year * Number of positive news articles		0.013
		(0.002)
Return previous year * Number of neutral news articles		-0.022
		(0.000)
Return previous year * Number of negative news articles		-0.035
		(0.000)
Sigma	0.118	0.120
	(0.007)	(0.008)
Rho	0.652	0.676
	(0.046)	(0.026)
Log likelihood	-5809	-5772
N	21052	21052

Table 24. Relation of Directional Fund Flows to Media Coverage considering Age of Fund for Non-Index Funds

The sample includes non-index funds with a growth objective that existed over the 1994-2000 time period. This table provides results from the Heckman maximum likelihood model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds in the sample during the period, the lagged flow to the fund for the previous period, the return to the fund for the previous year, log of the fund's total net assets under management (TNA) for the previous period, the fund's age in years, the fund's expense ratio for the previous year, and the volatility of fund return for the previous year.
The media coverage variables included in Model 1 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age variable and the number of news articles and between the age variable and the cumulative number of news articles. The media coverage variables included in Model 2 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age group dummy variables and the number of news articles and between the age group dummy variables and the cumulative number of news articles. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.176	-0.133
	(0.347)	(0.449)
Market Flow	0.260	0.260
	(0.010)	(0.010)
Flow (t-1)	0.001	-0.013
	(0.983)	(0.391)
Return previous year	0.090	0.084
	(0.000)	(0.000)
Log TNA (t-1)	0.005	0.003
	(0.524)	(0.621)
Age	0.001	0.001
	(0.026)	(0.166)
Expense ratio (t-1)	-0.151	-0.154
	(0.037)	(0.032)
Volatility of return previous year	0.281	0.123
	(0.301)	(0.626)

Number of news articles	0.006	0.002
	(0.000)	(0.110)
Cumulative number of news articles	-0.039	-0.026
	(0.000)	(0.002)
Age * Number of news articles	0.001	
	(0.002)	
Age * Cumulative number of news articles	0.001	
	(0.010)	
Young funds * Number of news articles		0.013
		(0.000)
Young funds * Cumulative number of news articles		-0.007
		(0.493)
Old funds * Number of news articles		-0.002
		(0.267)
Old funds * Cumulative number of news articles		0.007
		(0.362)
Sigma	0.116	0.105
	(0.006)	(0.002)
Rho	0.637	0.535
	(0.062)	(0.204)
Log likelihood	-5735	-5735
N	21052	21052

Table 25. Granger Causality Test Results

The sample includes growth funds that existed over the 1994-2000 time period. This table provides summary results for a vector autoregression (VAR) analysis of the relation between media coverage and fund flows. Fund flows are measured as percentage net monthly flow into the fund News is measured by the dummy variable for the existence of any news article portraying the fund in each period. This table provides summary results from three sets of Granger causality tests performed for each direction of causality. In addition to the zero-lag variable, Model 1 includes a single lag variable, Model 2 includes two lags of each variable and Model 3 includes three lags of each variable. Panel A presents the results for the regressions testing media coverage granger-causing mutual fund flows. Panel B presents the results for fund flows granger-causing media coverage of mutual funds. The table also provides the statistic for the Wald test for Granger causality for each regression. p-values are provided in parentheses below coefficient estimates.

Panel A : Does past media coverage granger-cause fund flows ?

	Model 1	**Model 2**	**Model 3**
	One Lag	**Two Lags**	**Three Lags**
Intercept	0.666	0.632	0.628
	(0.000)	(0.000)	(0.000)
Flow (t-1)	0.007	0.019	0.018
	(0.003)	(0.005)	(0.008)
Flow (t-2)		0.039	0.039
		(0.000)	(0.000)
Flow (t-3)			0.023
			(0.001)
News (t)	0.330	0.542	0.583
	(0.077)	(0.117)	(0.102)
News (t-1)	0.330	0.207	0.241
	(0.449)	(0.552)	(0.500)
News (t-2)		0.089	0.131
		(0.796)	(0.713)
News (t-3)			-0.189
			(0.592)

Wald test for Granger-Causality:			
Chi-Square	5.28	5.77	8.43
p-value	(0.072)	(0.449)	(0.751)

Panel B : Do past fund flows granger-cause media coverage ?

	Model 1 **One Lag**	**Model 2** **Two Lags**	**Model 3** **Three Lags**
Intercept	0.104	0.074	0.056
	(0.000)	(0.000)	(0.000)
News (t-1)	0.462	0.328	0.259
	(0.000)	(0.000)	(0.000)
News (t-2)		0.290	0.211
		(0.000)	(0.000)
News (t-3)			0.238
			(0.000)
Flow (t)	0.001	0.001	0.001
	(0.084)	(0.170)	(0.197)
Flow (t-1)	0.007	0.006	0.005
	(0.479)	(0.525)	(0.611)
Flow (t-2)		0.001	0.001
		(0.796)	(0.950)
Flow (t-3)			0.001
			(0.821)
Wald test for Granger-Causality:			
Chi-Square	0.88	2.75	4.30
p-value	(0.644)	(0.839)	(0.977)

Figure 1. Growth in Total Net Assets (in $ million)

The sample includes growth funds that existed over the 1994-2000 time period. This figure shows the growth in the total net assets under management for the sample funds through the sample period. The graph plots the average total net assets of all the funds in the sample for each month in the period.

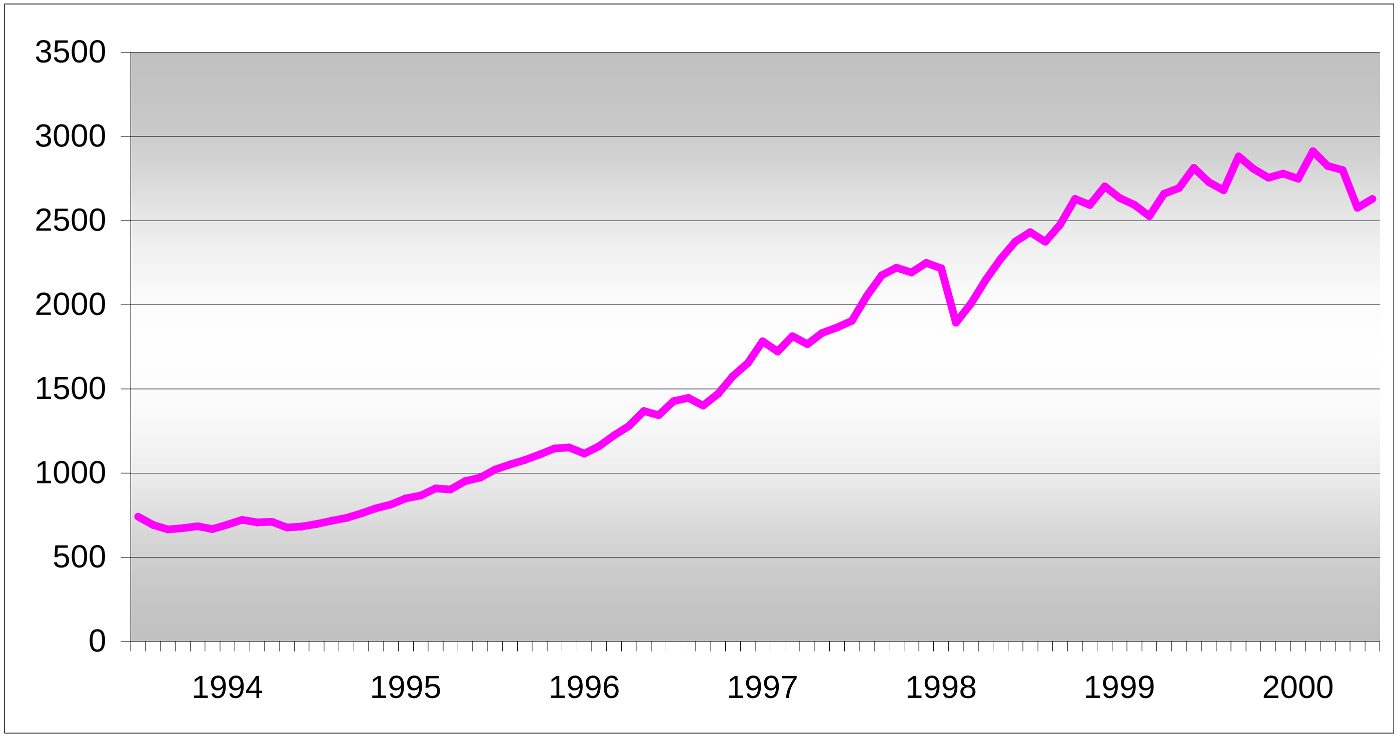


FIGURE 2. CHANGES IN PERCENTAGE FUND FLOWS

The sample includes growth funds that existed over the 1994-2000 time period. This figure shows the changes in the percentage net flows into the funds through the sample period. The fund flow for each fund is calculated as $Flow_{i,t}=\{TNA_{i,t}-(TNA_{i,t-1}*(1+R_{i,t}))\}/TNA_{i,t-1}$ where TNA is the total net assets for each fund and R is the monthly total return for each fund. The graph plots the average percentage fund flows in the sample for each month in the period.

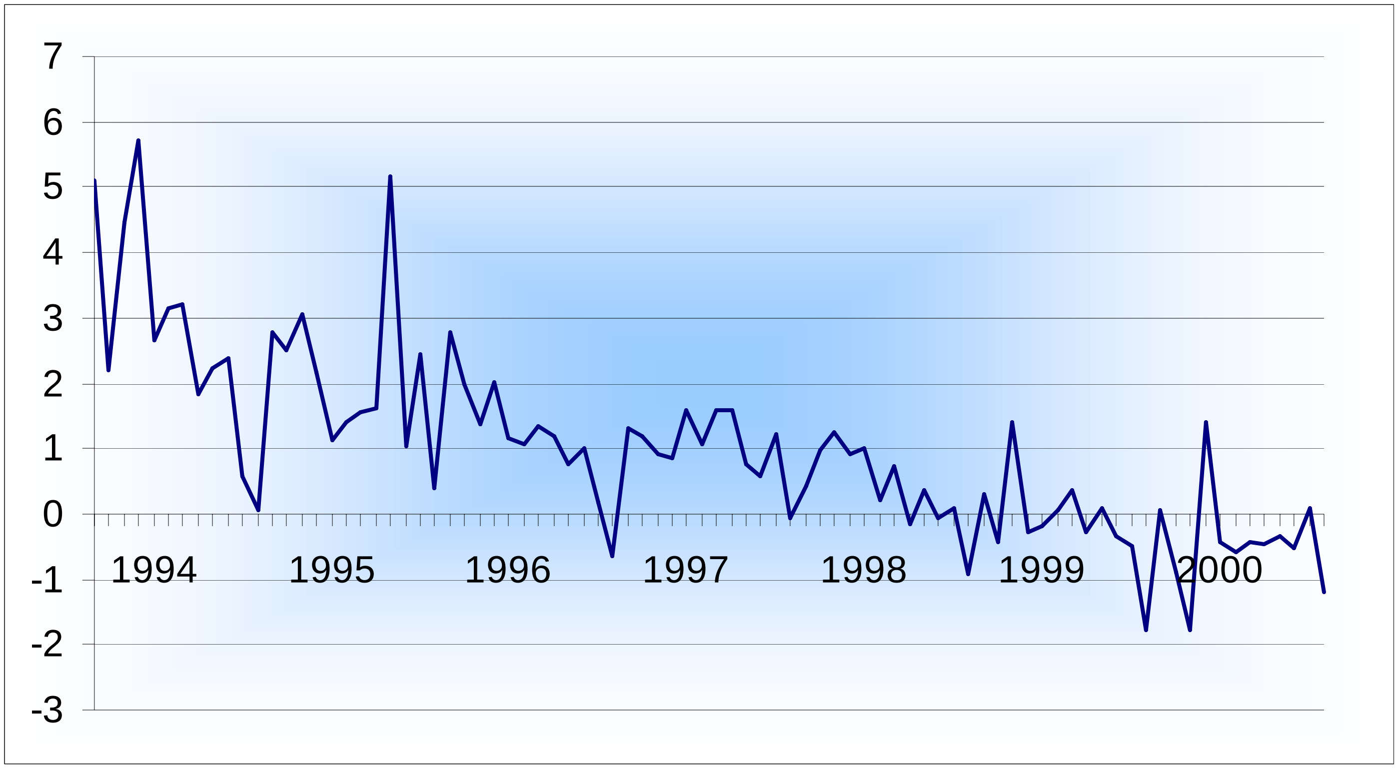


Figure 3. Monthly Return

The sample includes growth funds that existed over the 1994-2000 time period. This figure shows the changes in the monthly total return on the fund through the sample period. The graph plots the average total return for all the funds in the sample for each month in the period.

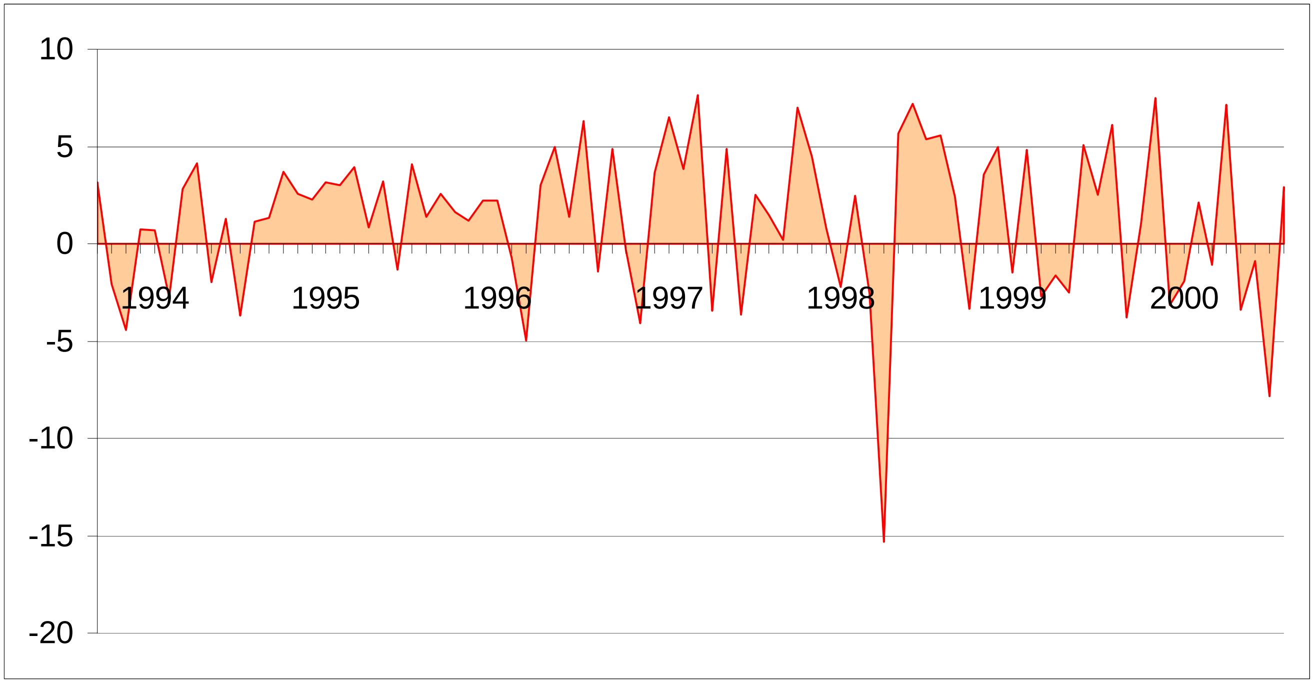


Figure 4A. Probability of Media Coverage

This figure shows the trend in the average probability of a fund receiving any media coverage through the sample period.

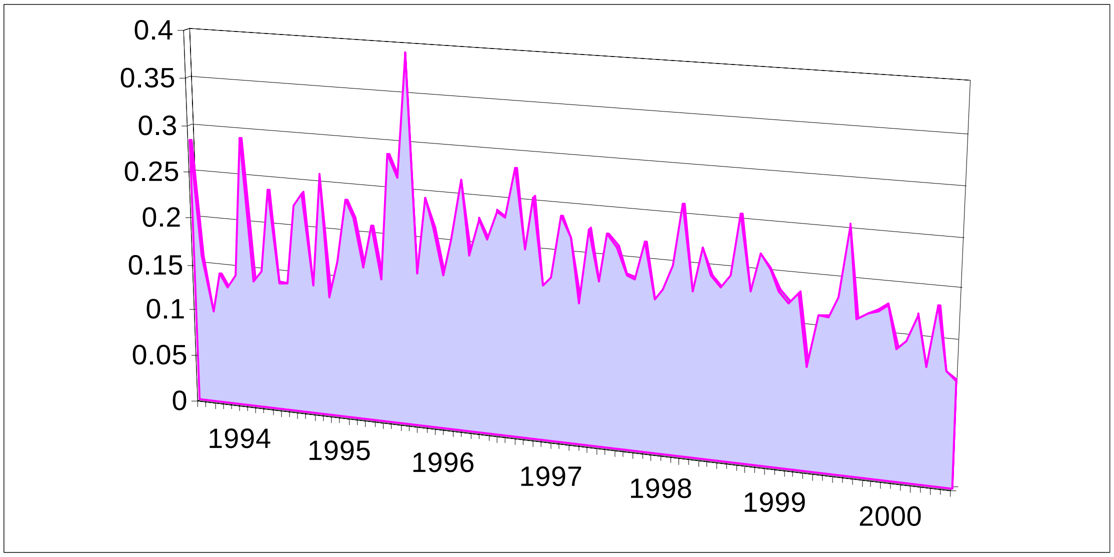


Figure 4B. Seasonality in Media Coverage

This figure shows the seasonality in the probability of a fund receiving any media coverage. The average probability of overall media coverage is calculated for the whole sample for each calendar month.

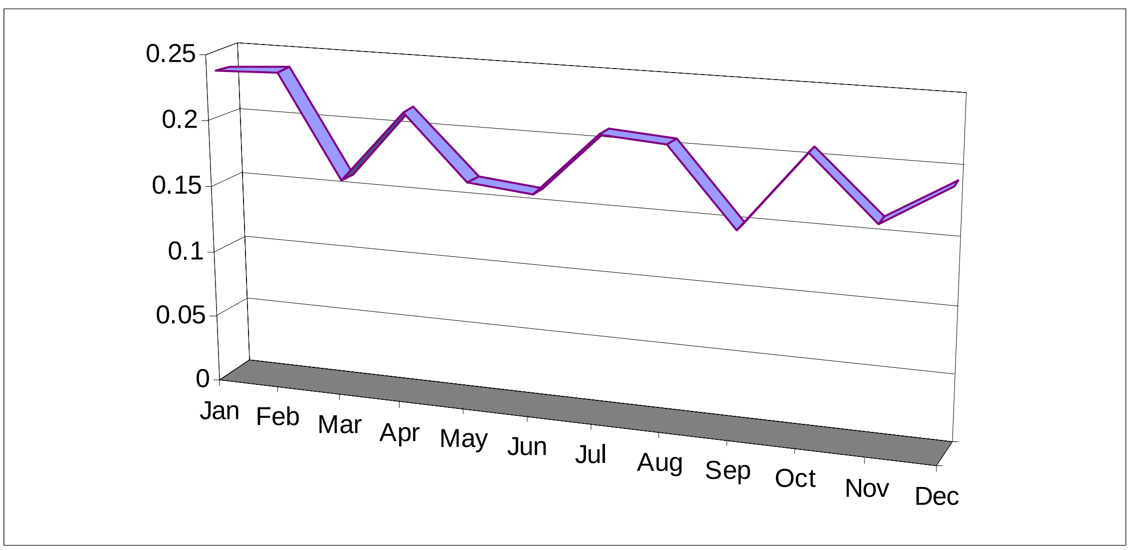
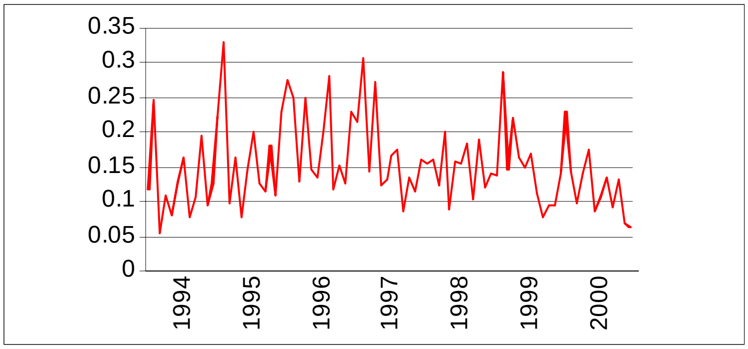


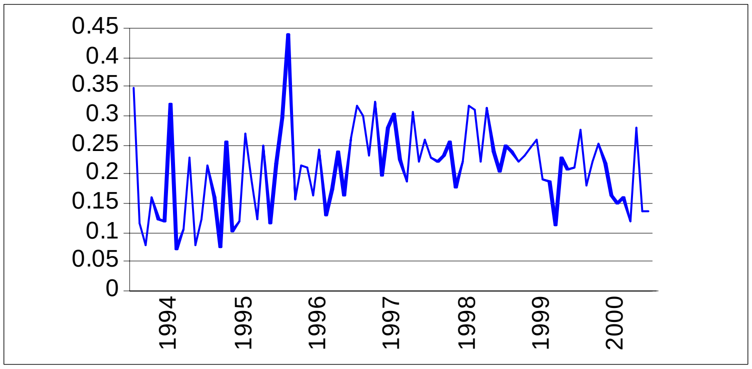
Figure 5. Posture of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. The figures below show the trend in the frequency of media coverage for funds in the sample through the sample period. The figures depict the trends for positive, neutral and negative media coverage. The average frequency for each posture of media coverage is calculated for the whole sample for each calendar month.

5A. Frequency of Positive Media Coverage



5B. Frequency of Neutral Media Coverage



5C. Frequency of Negative Media Coverage

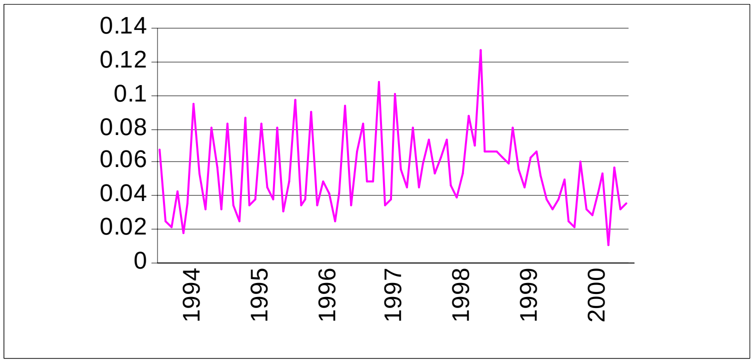
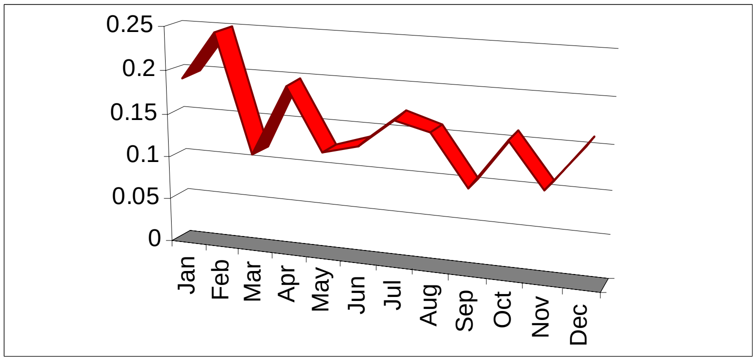


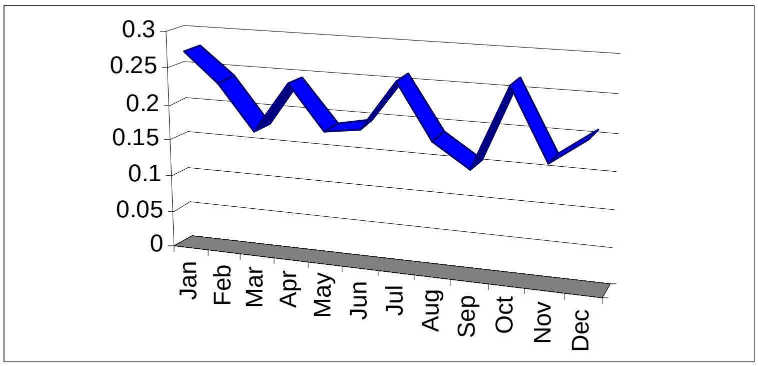
Figure 6. Seasonality in Posture of Media Coverage

The figures below show the seasonality in the frequency of media coverage for funds in the sample. Each figure focuses on a specific category of media coverage based on the posture of the news articles with regard to the fund portrayed. The average frequency of each category of media coverage is calculated for the whole sample for each calendar month.

6A. Frequency of Positive Media Coverage



6B. Frequency of Neutral Media Coverage



6C. Frequency of Negative Media Coverage

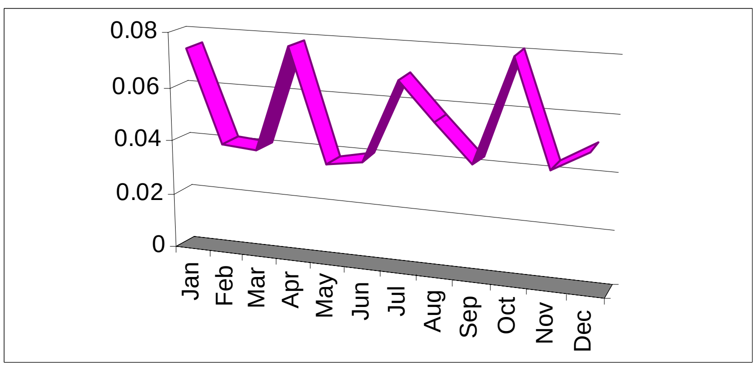
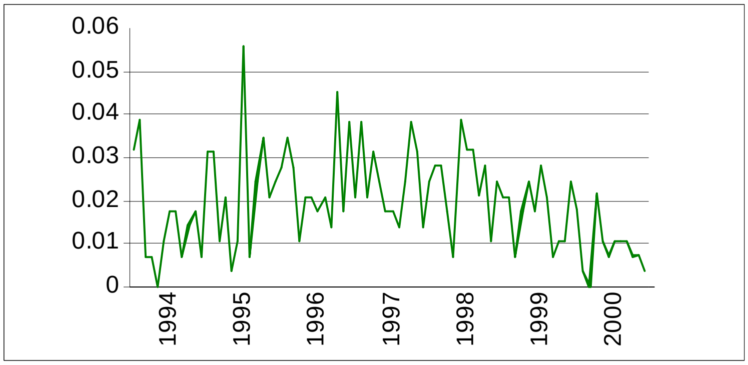


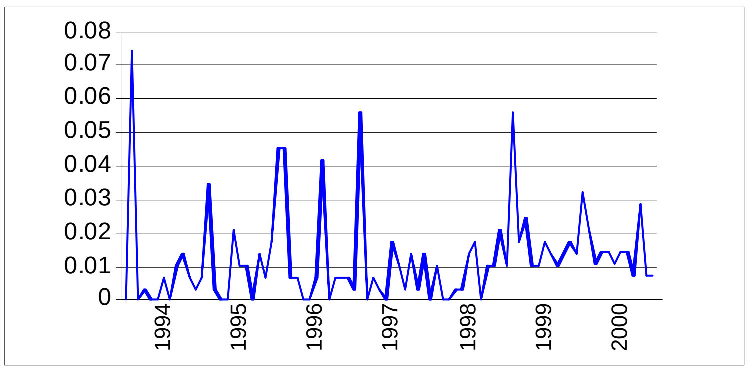
Figure 7. Type of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. The figures below show the trend in the frequency of media coverage for funds in the sample through the sample period. Each figure focuses on a specific type of media coverage as specified in the section on the classification of news articles. The average frequency of each type of media coverage is calculated for the whole sample for each calendar month.

7A. Frequency of 'Performance' Type Media Coverage



7B. Frequency of 'Ranking' Type Media Coverage



7C. Frequency of 'Mention' Type Media Coverage

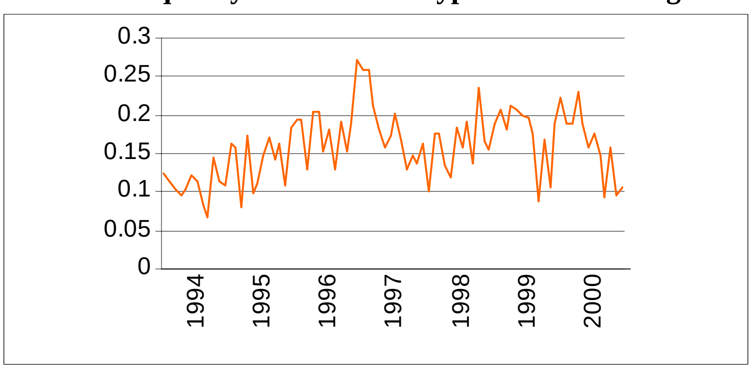
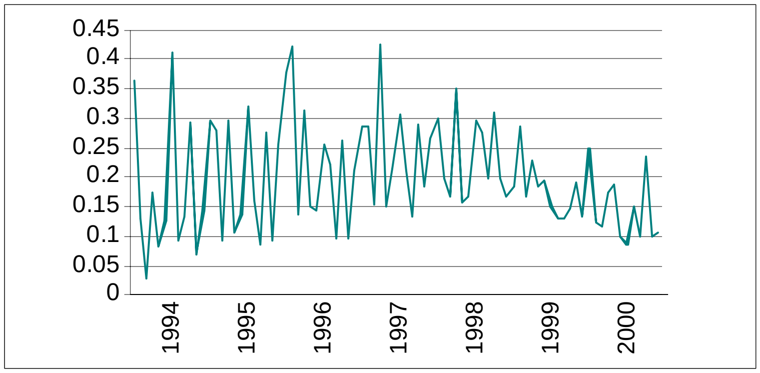


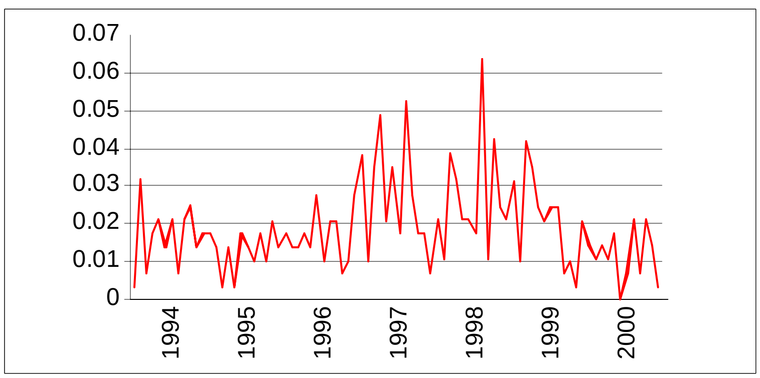
Figure 7. Type of Media Coverage – Continued

The sample includes growth funds that existed over the 1994-2000 time period. The figures below show the trend in the frequency of media coverage for funds in the sample through the sample period. Each figure focuses on a specific type of media coverage. The average frequency of each type of media coverage is calculated for the whole sample for each calendar month.

7D. Frequency of 'Tables' Type Media Coverage



7E. Frequency of 'Feature' Type Media Coverage



7F. Frequency of 'Interview' Type Media Coverage

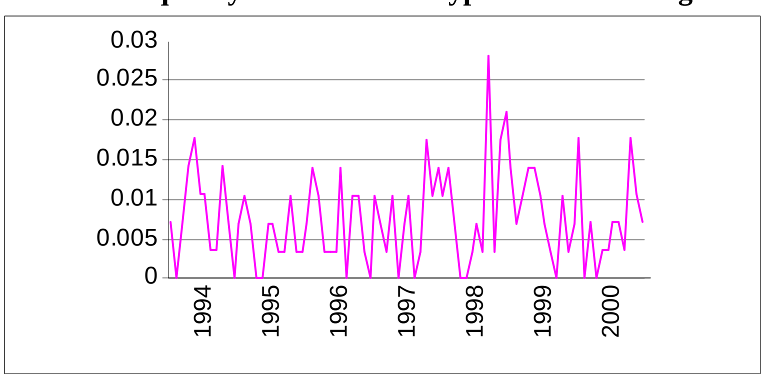
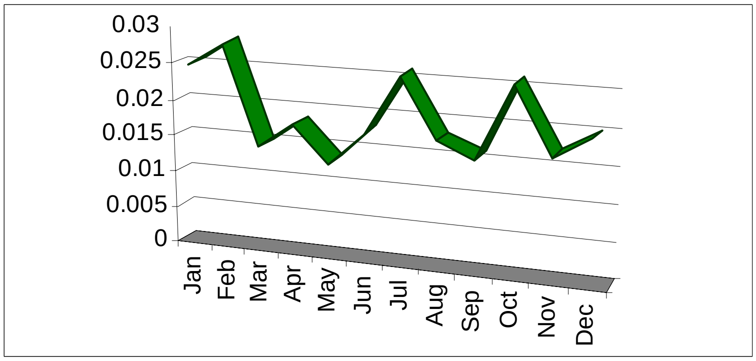


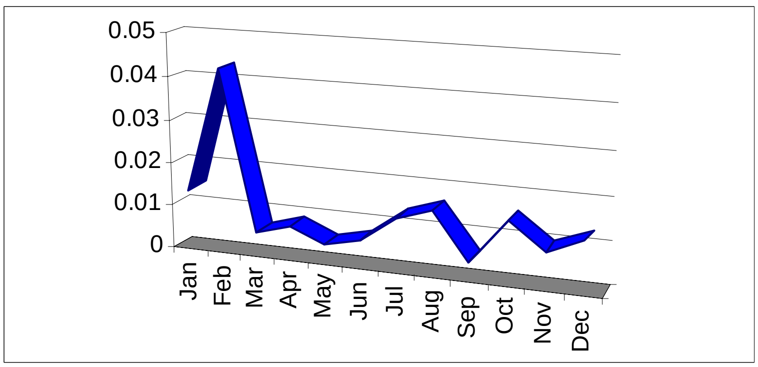
Figure 8. Seasonality in Type of Media Coverage

The figures below show the seasonality in the frequency of media coverage for funds in the sample. Each figure focuses on a specific type of media coverage. The average frequency of each type of media coverage is calculated for the whole sample for each calendar month.

8A. Frequency of 'Performance' Type Media Coverage



8B. Frequency of 'Ranking' Type Media Coverage



8C. Frequency of 'Mention' Type Media Coverage

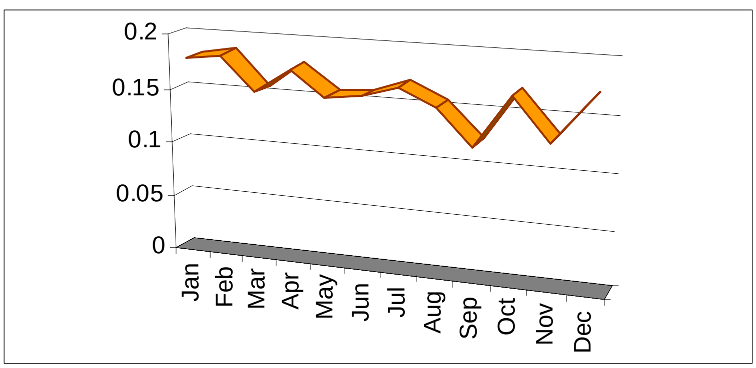
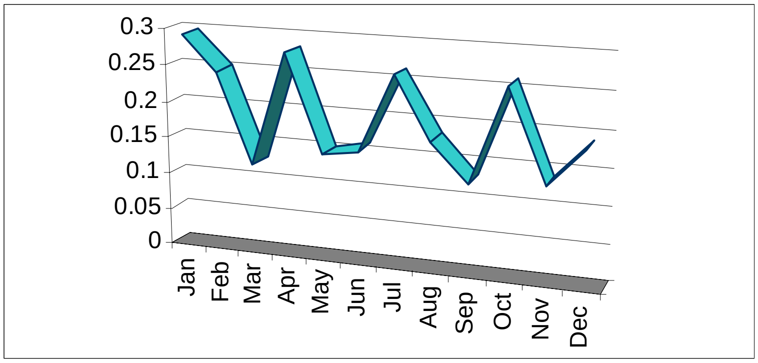


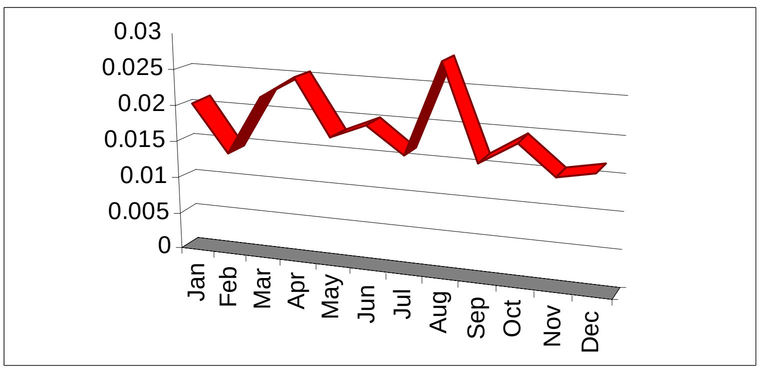
Figure 8. Seasonality in Type of Media Coverage – Continued

The figures below show the seasonality in the frequency of media coverage for funds in the sample. Each figure focuses on a specific type of media coverage. The average frequency of each type of media coverage is calculated for the whole sample for each calendar month.

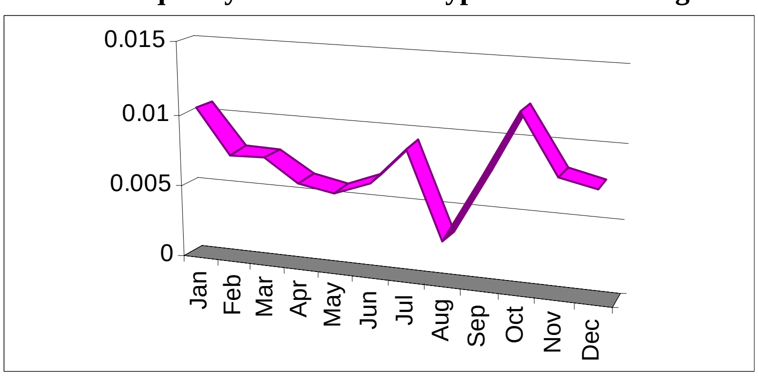
8D. Frequency of 'Tables' Type Media Coverage



8E. Frequency of 'Feature' Type Media Coverage



8F. Frequency of 'Interview' Type Media Coverage



Appendix A

Mutual Fund Share Classes

This appendix provides a background on mutual fund share classes.[30] Mutual funds are often classified according to the class of shares that fund sponsors offer to investors: primarily load or no-load classes. Load classes generally serve investors who hold funds through financial advisers; no-load fund classes usually serve investors who purchase funds without the assistance of a financial adviser or who choose to compensate the financial adviser separately. More than half of all mutual funds offer two or more share classes. Funds that sell through financial advisers offer more than one share class to provide investors with several ways to pay for the services of financial advisers.

Load Share Classes

Load share classes - typically labeled class A, B, and C shares - usually include a sales load and/or a 12b-1 fee. The sales load and 12b-1 fees are used to compensate financial advisers for their services.

Class A shares

Class A shares represent the traditional means of paying for investment advice and assistance. Class A shares generally charge a front-end sales load at the time of the purchase as a percentage of the sales price or offering price. This share class also often has a 12b-1 fee of about 0.25 percent. Class A shares are sometimes used in employer-

[30] Source: Investment Company Institute, 2006

sponsored retirement plans, and funds usually waive the front-end sales load for these investors.

Class B shares

Class B shares typically do not have a front-end sales load. Investors using B shares pay for financial advisers through a combination of an annual 12b-1 fee, usually 1 percent, and a contingent deferred sales load (CDSL). The CDSL is triggered if fund shares are redeemed before a fixed number of years of ownership. The CDSL decreases the longer the investor owns the shares and reaches zero typically after shares have been held six or seven years. After six to eight years, B shares usually convert to A shares, which have a lower 12b-1 fee.

Class C shares

Class C shares generally do not have a front-end load. Investors in this share class compensate financial advisers with a combination of an annual 1 percent 12b-1 fee and a 1 percent CDSL paid directly by shareholders if they sell their shares within the first year after purchase. This share class, unlike B shares, typically does not convert to A shares.

NO-LOAD SHARE CLASSES

No-load share classes have no front-end load or CDSL and have a 12b-1 fee of 0.25 percent or less. Originally, no-load share classes were offered by mutual fund sponsors that sold directly to investors. Now, however, investors can purchase no-load funds through employer-sponsored retirement plans, mutual fund supermarkets, discount brokerage firms, and bank trust departments. Some financial advisers who charge investors separately for their services rather than through a load or 12b-1 fee also use no-load share classes.

Appendix B

Classification of News Articles

I trained and supervised a team of 6 research assistants to help in the collection and classification of part of the news database. I individually trained and monitored each member of the team for a significant period of time. After each member was fully trained, I assigned funds for which they were to search and classify news articles. I picked a random sample of 20% of the articles classified by each of them and reclassified them to ensure total compliance with the training given.

Methodology of Classification

The goals were:

- To search for mentions of each of the 286 sample funds in 12 major publications through the Dow Jones Retrieval Service (now called Factiva) for the 1994-2000 sample period.
- To subsequently classify each of the 9,984 news articles in terms of a given set of characteristics.

The search criteria were as follows for each of the 286 funds:

- The following list of publications was selected:

  Boston Globe, *Barron's*, *Business Week*, *Forbes*, *Fortune*, *Los Angeles Times*, *Money*, *New York Times*, *US News & World Report*, *USA Today*, *Washington Post*, and *Wall Street Journal*.

- The date range was 01/01/1994 – 12/31/2000.
- The fund name was entered as the search phrase.

Each article that was brought up in the search was then read and classified. The classification and the corresponding characteristics of the news articles were then entered in a spreadsheet. The following details were entered in the spreadsheet for each article for every fund:

- Fund name used as the search phrase
- Date of the article
- Name of the publication in which the article appeared
- Page number, if available, on which the article appeared
- Code for the tone of the article as classified into one of the specified categories:
  - PS: Positive

    The fund is portrayed in a clearly positive light. There is a substantial emphasis on the positive aspects of the fund.
  - NP: Neutral-Positive

    The fund is portrayed in a marginally positive way. Bulk of the news story is not dedicated to portraying the fund in a positive light but there is still a positive posture.
  - NT: Neutral

    The article mentions the fund in a neutral way with no positive or negative slant towards the fund.
  - NN: Neutral-Negative

The fund is portrayed in a marginally negative way. Bulk of the news story is not dedicated to portraying the fund in a negative light but there is still a negative posture.

- NG: Negative

  The fund is portrayed in a clearly negative light.

- Code for the type of the article as classified into one of the specified categories:
  - FR: Feature

    The fund is portrayed in some detail.
  - IN: Interview

    The article portrays an interview, usually with a fund manager.
  - RK: Ranking

    The fund is portrayed with a numerical ranking.
  - PF: Performance

    The article portrays the fund's performance.
  - MN: Mention

    The article merely mentions the fund.
  - TB: Tables

    The fund is portrayed only in a table.

The criteria for classification of the type of news articles are as follows:

The type of news article is classified in the order of the above listing. First, one checks if the article can be classified as FR. If an article focuses on a fund to the extent that a considerable portion, if not all, of the article portrays the fund, then the article is classified as FR. If an article cannot be classified as FR, then one sees of it can be classified as IN. If an article details an interview, usually with a fund manager, then it is classified as IN.

If an article cannot be classified as FR or IN, then the next step is to look for a numerical ranking for the fund in the article. If the fund is assigned a numerical ranking anywhere in the text or table section of the news story, then RK is assigned to the story. If the article discusses the fund's performance but does not assign a numerical ranking, then the article is classified as PF.

If a news story does not qualify for FR, IN, RK or PF, and merely mentions the fund in a non-performance context, then the news story is classified as MN. When the fund is mentioned nowhere in the text part of the article, and only appears in a table (without a numerical ranking), then the article is classified as TB.

Computer program versus manual classification

Computer programs used for quantitative content analysis are usually pure word count programs. Such programs typically gather information by counting the words that fall within certain predetermined categories, for example, 'positive' or 'negative'. The advantage to this method is that it can be perfectly replicated, but there are a lot of disadvantages as well. The word categories are neither mutually exclusive nor exhaustive. The programs do not usually categorize combinations of words that often possess different meanings from the constituent words. Tetlock (forthcoming JF) provides the following example of this fault. Consider the sentences: "No, the economy is not strong" and "It is not that the economy is not strong." The computer program understands and categorizes all of the important words in both sentences, but pure category counts would suggest that these sentences have identical meanings. In fact, the sentences have opposite meanings.

All of the above disadvantages can potentially be eliminated by adopting the human or manual approach to news classification. Each researcher is specifically trained to read through the entire article and then subsequently classify it based on their understanding of its content. This ensures that the semantic and stylistic noise does not in any way obscure the interpretations of the news stories. Individually monitored training also ensures consistency when repeated by different individuals. Finally, reclassification by the team leader of a random portion of the news stories that are classified by the research team ensures total compliance.

Construction of news variables based on classification

I build on the classification of tone and type of news article to construct the other news variables as follows:

News Flag (dummy) variables:

Article Flag:

$\mathrm{ARTF}_{i,t}$ = 1 if fund i has a news article in period t, else = 0

Positive Flag:

$\mathrm{PBF}_{i,t}$ = 1 if fund i has a PS or NP news article in period t, else = 0 Neutral Flag:

$\mathrm{UBF}_{i,t}$ = 1 if fund i has a NT news article in period t, else = 0

Negative Flag:

$\mathrm{NBF}_{i,t}$ = 1 if fund i has a NG or NN news article in period t, else = 0

News Count variables:

Total Article count:

$$ARTC_{i,t} = \#PS_{i,t} + \#NP_{i,t} + \#NT_{i,t} + \#NN_{i,t} + \#NG_{i,t}$$

Positive Count:

$$PBC_{i,t} = \#PS_{i,t} + \#NP_{i,t}$$

Neutral Count:

$$UBC_{i,t} = \#NT_{i,t}$$

Negative Count:

$$NBC_{i,t} = \#NG_{i,t} + \#NN_{i,t}$$

Feature and Interview Count:

$$FRIN_{i,t} = \#FR_{i,t} + \#IN_{i,t}$$

Mention and Tables Count:

$$MNTB_{i,t} = \#MN_{i,t} + \#TB_{i,t}$$

Performance and Ranking Count:

$$PFRK_{i,t} = \#PF_{i,t} + \#RK_{i,t}$$

Cumulative Article Count:

$$cumARTC_{i,t} = \sum_{t=1}^{n} ARTC_{i,t} \text{, where}$$

t = 1 is the first period in the sample (Jan 1994)

t = n is the current period

Circulation news variable:

$$\text{Circ}_{i,t} = \text{A}_{j,i,t} * \frac{\text{C}_j}{\sum_{j=1}^{12} \text{C}_j} \quad \text{, where}$$

$A_{j,i,t}$ = # articles in periodical *j* for fund *i* in period *t*

C_j = circulation of periodical *j*

Appendix C

Derivation of the Heckman Model

The basic idea of sample selection models is that the outcome variable, y, is only observed if some criterion, defined with respect to a different variable, z, is met. The standard form of the model has two stages:

- In the first stage, a dichotomous variable z (= 0 or 1) determines whether or not y is observed, y being observed only if z = 1.
- In the second stage, the expected value of y is modeled, conditional on its having been observed.

The standard regression model $y_i = x'_i\boldsymbol{\beta} + \varepsilon_i$ assumes that the data represents a random sample from some population of interest. When, for some reason, the sample is non-random, the OLS estimator can produce biased and inconsistent estimates.

The Heckman model can be defined by a set of two equations that describe the selection regression model (probit) and the outcome regression model (linear).

Selection Regression Model:

$$z_i^* = w_i'\gamma + u_i \tag{C.1}$$

$$z_i = \begin{cases} 1 & \text{if } z_i^* > 0 \\ 0 & \text{if } z_i^* \le 0 \end{cases}$$

Outcome Regression Model:

$$y_i^* = x_i'\beta + \varepsilon_i \qquad \text{(C.2)}$$

$$y_i = y_i^* \text{ if } z_i = 1$$

$$y_i \text{ not observed if } z_i = 0$$

where u_i and ε_i have a joint distribution that is bivariate normal with zero mean, standard deviations of σ_u and σ_ε, and correlation of ρ. *z* is the variable that the selection is based on and *y* is observed when *z* has a value of 1. The observed dummy variable z is a realization of an unobserved continuous variable z^*. The latent variable z^* can be interpreted as the propensity to be included in the sample.

The probit model for the probability of z = 1 is estimated using all the observations and yielding coefficient vector α:

$$pr\ (z_i = 1) = \Phi\ (w_i'\gamma)$$

Since α and σ_u are not separately identifiable in the probit, we may without loss of generality normalize u such that its variance is equal to 1, giving us $\sigma_u = 1$. The second step is to estimate the expected value of y, conditional on z = 1 and on the vector x_i.

$$E\ [y_i \mid z = 1, x_i] = x_i'\beta + E\ [\varepsilon_i \mid z_i = 1] \qquad \text{(C.3a)}$$

$$= x_i'\beta + E\ [\varepsilon_i \mid u_i > w_i'\gamma]$$

To evaluate the conditional expectation of ε in equation (C.3a), one can use the result from statistical theory that says that the expected value of one of the variables in a bivariate distribution (in this case, ε) censored with respect to the value of the other variable (in this case, u) is given by

$$E\,(\varepsilon_i \mid u_i > w_i'\gamma) = \rho\sigma_u\sigma_\varepsilon \frac{\phi(w_i'\gamma)}{\Phi(w_i'\gamma)} \tag{C.3b}$$

where ϕ and Φ are the standardized normal density and distribution functions respectively.

Inserting equation (C.3b) into equation (C.3a),

$$E\,(y_i \mid z_i = 1, x_i) = x_i'\beta + \rho\sigma_u\sigma_\varepsilon \frac{\phi(w_i'\gamma)}{\Phi(w_i'\gamma)} \tag{C.3c}$$

Heckman's two-step procedure consists of first estimating the selection equation (C.1) as a probit model. Using the probit results, the estimate of ϕ_i/Φ_i (the inverse Mill's ratio or Hazard ratio, symbolized by λ_i) is computed for the subsample where z = 1. Then, for this same subsample, OLS is used to regress y on x_i with the estimate of λ_i as an additional explanatory variable:

$$E\,(y_i \mid z_i = 1, x_i) = x_i'\beta + \theta\hat{\lambda}_i \tag{C.4}$$

This will yield estimates of β and θ. θ is an estimate of ρ times σ_ε, which, because $\sigma_u = 1$ is equal to the covariance between u and ε ($\sigma_{u\varepsilon}$),

$$\theta = \rho\sigma_{\varepsilon} = \frac{\sigma_{u\varepsilon}}{\sigma_u\sigma_{\varepsilon}}\sigma_{\varepsilon} = \sigma_{u\varepsilon}$$

Heckman (1979) shows that selection bias can be thought of as a form of omitted variable bias. The two-step model explicitly addresses bias caused by correlation of the regressor with omitted variables, by adding a term to the least squares regression that represents the non-zero expectation of the error term. T'o simply take those cases with observations on y and regress y_i on x_i, will result in biased estimates of the vector β. Equation (C.4) shows that it is the omitted variable λ that causes the OLS estimation to be biased, where λ is the *Hazard ratio*, also known as *Inverse Mill's ratio*. The problem of sample-selection bias thus becomes equivalent to a misspecification problem arising through the omission of a regressor variable. However, the OLS estimates of β will be unbiased if there is no correlation between the error terms of the two equations. If $\rho = 0$, this means that θ in Equation (C.4) must also be zero, and thus Equation (C.4) reduces to the usual OLS equation. This corresponds to the situation in which selection and outcome are independent.

Heckman's two-stage estimator is perhaps the most widely used approach to selection bias. Instead of using maximum likelihood, the two-stage method permits estimation of the model by use of simpler statistical procedures like least squares and probit analysis. This is particularly useful if maximum likelihood will not converge. However, the Heckman two-step method gives consistent estimates of β but not of σ, nor are the asymptotic standard errors consistent. As a result, adjustments must be made to the regression results. These adjustments are relatively straightforward and are provided below.

Define

$$\delta_i = -\lambda_i (z_i + \lambda_i)$$

where $z_i = w_i'\gamma$ from the probit step.

Then, let s_ε be the uncorrected estimate of σ_ε from the regression (second) step of the Heckman procedure, and let S be the sum of squared deviations from this regression:

$$S = \sum_{i \in \{z_i = 1\}} \left(y_i - \hat{y}_i \right)^2$$

where the sum is taken over all observations for which z = 1. The correct asymptotic estimate of σ_ε is given by

$$\hat{\sigma}_\varepsilon = \frac{1}{N} \left[S - \hat{\theta}^2 \sum_{i=1}^{N} \delta_i \right]^{\frac{1}{2}} \tag{C.5}$$

where N is the number of observations for which z = 1 and θ is the estimated regression coefficient for λ.[31]

Heteroscedasticity in model (C.4) causes the standard errors to be incorrect. The standard errors are incorrect also because the use of an estimate of λ, rather than λ itself, means that the standard errors for the β coefficients will need to take account of the error in the estimate of λ. The uncorrected OLS standard errors can be either larger or smaller than the corrected ones. Hence, they cannot be used as lower bounds on the true standard errors. The formula for the correct covariance matrix of the estimates of β and σ_ε, V, is given by the matrix equation

[31] Heckman (1979), Greene (2000).

$$V = \sigma_\varepsilon^2 \left(X^{*'} X^* \right)^{-1} \left[X^{*'} \left(I - \rho^2 \Delta \right) X^* + \rho^2 \left(X^* \Delta W \right) \sum \left(W' \Delta X^* \right) \right] \left(X^{*'} X^* \right)^{-1} \tag{C.6}$$

The standard errors of the parameter estimates are given by the square root of the diagonal elements of V. Here, X^* is the matrix [x:λ]; W is the matrix of explanatory variables from the probit equation; Δ is the matrix with the elements δ_i on its diagonal, zeros elsewhere; I is the identity matrix; and Σ is the asymptotic covariance matrix for the parameters of the probit equation. ρ is estimated by

$$\hat{\rho} = \frac{\hat{\theta}}{\hat{\sigma}_\varepsilon} \tag{C.7}$$

An alternate approach to estimate the Heckman model is Maximum likelihood estimation, which overcomes the difficulties associated with the two-step approach – notably, that the estimated standard errors of the coefficients and the estimate of sigma are all incorrect. The log-likelihood function needs to be specified before estimation of the model.

All those cases where z = 0 contribute $1 - \Phi\,(w_i'\gamma)$ to the likelihood.

Those cases where z = 1 contribute

$$\Phi\,(w_i'\gamma) \times \frac{1}{\sigma} \phi\,(y_i \mid z_i = 1) \tag{C.8}$$

where σ is the standard deviation of y^* conditional on $z = 1$ and $\phi(y_i \mid z_i = 1)$ is the conditional density function of y^* given $z = 1$. Equation (C.8) is the expression for the probability of being selected multiplied by the density of y conditional on having been selected. To put this into a tractable form for estimation, some further manipulations are needed. Amemiya (1984) demonstrates that equation (C.8) can be written as

$$\Phi\left[\frac{w_i'\gamma + \rho\left(\frac{y_i - x_i'\beta}{\sigma_\varepsilon}\right)}{\left(1-\rho^2\right)^{\frac{1}{2}}}\right] \times \frac{1}{\sigma_\varepsilon}\,\Phi\left(\frac{y_i - x_i'\beta}{\sigma_\varepsilon}\right) \tag{C.9}$$

The log-likelihood can be derived by inserting the above expression for those observations where z = 0 and taking logarithms. The log-likelihood function of the Heckman selection model is written as:

$$\ln L = \sum_{i\in\{z_i=0\}} \ln\left[1-\Phi(w_i'\gamma)\right] + \sum_{i\in\{z_i=1\}} \ln\left(\frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}}\right)$$

$$- \sum_{i\in\{z_i=1\}} \frac{1}{2\sigma_\varepsilon^2}\left(y_i - x_i'\beta\right)^2 + \sum_{i\in\{z_i=1\}} \ln\Phi\left[\frac{w_i'\gamma + \rho\left(\frac{y_i - x_i'\beta}{\sigma_\varepsilon}\right)}{\left(1-\rho^2\right)^{\frac{1}{2}}}\right]$$

(C.10)

Once the likelihood function is specified, nonlinear optimization is used to maximize the likelihood function and calculate the standard errors of the estimates. The likelihood equation can be used to show the circumstances under which this model reduces to something simpler. If $\rho = 0$, then equation (C.10) can be split into two parts:

- The first part is a probit for the probability of being selected and comprises of the first and fourth terms in the equation.
- The second part is an OLS regression for the expected value of y in the selected subsample and comprises of the second and third terms in the equation.

Furthermore, because these two parts share no common parameters, they can be estimated separately. This shows that if there is no residual correlation between ε and u, the simple OLS approach is adequate. Therefore, it is not the fact that observations on y are only available for a selected sample that causes the difficulty; rather, it is that this selection is not random with respect to y. This is particularly applicable in the case of my research on media coverage of mutual funds.

The maximum likelihood estimates have the desirable properties of being consistent and asymptotically efficient. However, the main criticism of the maximum likelihood estimation method is that it is extremely sensitive to the correct specification of the underlying distribution. On the other hand, Winship and Mare (1992) note that while Heckman’s two-stage estimator is the most widely used approach to selection bias, its results may be sensitive to violations of its assumptions about the way selection occurs.

Breen (1996) recommends that for the sample-selection model, one should first compute the inverse Mill's ratio instrument using a probit model and regress this on the explanatory (x) variables that appear in the outcome equation. If the R^2 from this regression is around zero, then the outcome equation can be estimated using uncorrected OLS. If this is not so, then the two stages of the model should be estimated using maximum likelihood. If, however, the estimate of ρ in the maximum likelihood model is around zero, it is advisable once again to estimate the outcome equation as an uncorrected OLS regression. This is because the OLS estimates of β will differ little, if at all, from the maximum likelihood estimates but they will have smaller variances, particularly if R^2 is large. Hence, Breen (1996) suggests that when estimating the sample-selection model, one should use an uncorrected OLS for the outcome equation if either R^2 or the estimate of ρ is near zero; otherwise, the maximum likelihood should be used. I follow this recommendation from Breen.

I use the Heckman two-step model to first compute the Hazard ratio or the Inverse Mill's ratio (λ) using a probit model on the full sample and then regress this on the explanatory variables that appear in the outcome equation. I use the maximum likelihood model to estimate the correlation coefficient (ρ). I find that neither the R^2 from the earlier model nor the ρ from the latter model is around zero, thereby ruling out the presentation of uncorrected OLS estimates. Therefore, I present results from the Heckman maximum likelihood model for the empirical analysis of the media coverage of mutual funds in Chapter 6 of the dissertation. I present the corresponding results from Heckman two-step model in Appendix D.

Appendix D

Empirical Results from the Heckman Two-Step Model

Table D1. Relation of Fund Flows to Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman two-step model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds during the period, the lagged flow to the fund, log of the fund's total net assets under management (TNA) for the previous period, the return to the fund for the previous month, the return to the fund for the previous year, the fund's age in years, and the fund's expense ratio for the previous year.
The media coverage variable included in Model 1 is the actual number of news stories portraying the fund each month. The media coverage variable included in Model 2 is the actual number of news stories portraying the fund each month, weighted by the proportion of the circulation of each periodical to the total circulation of all the twelve periodicals. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.031	-0.035
	(0.345)	(0.292)
Market Flow	0.209	0.207
	(0.000)	(0.000)
Flow (t-1)	0.009	0.010
	(0.725)	(0.712)
Return previous month	-0.029	-0.028
	(0.000)	(0.000)
Return previous year	0.094	0.094
	(0.007)	(0.008)
Log TNA (t-1)	-0.088	-0.050
	(0.950)	(0.972)

Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.143	-0.128
	(0.037)	(0.060)
Number of news articles	0.002	
	(0.000)	
Circulation-weighted number of news articles		0.022
		(0.001)
Lambda	0.024	0.026
	(0.126)	(0.108)
R^2 (%)	6.78	6.71
N	4537	4537

Table D2. Relation of Fund Flows to Posture of Media Coverage

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman two-step model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds during the period, the lagged flow to the fund, log of the fund's total net assets under management (TNA) for the previous period, the return to the fund for the previous month, the return to the fund for the previous year, the fund's age in years, and the fund's expense ratio for the previous year.
The media coverage variables included in Model 1 are the dummy variables for the existence of news stories in each period, classified by the posture of the article towards the fund portrayed. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.018	-0.014
	(0.589)	(0.666)
Market Flow	0.198	0.195
	(0.000)	(0.000)
Flow (t-1)	0.003	0.005
	(0.904)	(0.863)
Return previous month	-0.033	-0.034
	(0.000)	(0.000)
Return previous year	0.084	0.084
	(0.017)	(0.018)
Log TNA (t-1)	-0.191	-0.196
	(0.892)	(0.890)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.093	-0.096
	(0.172)	(0.167)
Dummy for whether the fund has:		
Positive media coverage	0.014	
	(0.000)	
Neutral media coverage	0.007	
	(0.051)	

Negative media coverage	-0.010	
	(0.007)	
Number of positive news articles		0.006
		(0.000)
Number of neutral news articles		0.002
		(0.033)
Number of negative news articles		-0.006
		(0.006)
Lambda	0.018	0.018
	(0.262)	(0.250)
R^2 (%)	7.34	7.28
N	4537	4537

Table D3. Relation of Fund Flows to Interaction of Media Coverage and Fund Size

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman two-step model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds during the period, the lagged flow to the fund, log of the fund's total net assets under management (TNA) for the previous period, the return to the fund for the previous month, the return to the fund for the previous year, the fund's age in years, and the fund's expense ratio for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the log of the fund's TNA for the previous period and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the log of the fund's TNA for the previous period. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.066	-0.062
	(0.049)	(0.065)
Market Flow	0.209	0.187
	(0.000)	(0.000)
Flow (t-1)	0.004	-0.004
	(0.875)	(0.874)
Return previous month	-0.021	-0.026
	(0.002)	(0.000)
Return previous year	0.095	0.078
	(0.007)	(0.023)
Log TNA (t-1)	0.261	0.269
	(0.852)	(0.848)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.190	-0.097
	(0.006)	(0.172)
Number of news articles	0 .014	
	(0.000)	
Log TNA (t-1) * Number of news articles	-0.001	
	(0.000)	

Number of positive news articles		0 .031
		(0.000)
Number of neutral news articles		0.014
		(0.000)
Number of negative news articles		0 .003
		(0.339)
Log TNA (t-1) * Number of positive news articles		-0.003
		(0.000)
Log TNA (t-1) * Number of neutral news articles		-0 .001
		(0.000)
Log TNA (t-1) * Number of negative news articles		-0.001
		(0.084)
Lambda	0.030	0.028
	(0.056)	(0.079)
R^2 (%)	8.60	8.37
N	4537	4537

Table D4. Relation of Fund Flows to Interaction of Media Coverage and Fund Performance

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman two-step model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds during the period, the lagged flow to the fund, log of the fund's total net assets under management (TNA) for the previous period, the return to the fund for the previous month, the return to the fund for the previous year, the fund's age in years, and the fund's expense ratio for the previous year.
The media coverage variables included in Model 1 are the actual number of news stories portraying the fund each month, and the interaction terms between the return to the fund for the previous year and the number of news articles in the current period. The media coverage variables included in Model 2 are the numbers of news stories in each period, classified by the posture of the article towards the fund portrayed, and the interaction terms between these news count variables and the return to the fund for the previous year. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	-0.026	-0.029
	(0.437)	(0.378)
Market Flow	0.219	0.217
	(0.000)	(0.000)
Flow (t-1)	0.009	-0.002
	(0.731)	(0.928)
Return previous month	-0.034	-0.034
	(0.000)	(0.000)
Return previous year	0.113	0.113
	(0.001)	(0.001)
Log TNA (t-1)	-0.145	-0.125
	(0.918)	(0.929)
Age	-0.001	-0.001
	(0.000)	(0.000)
Expense ratio (t-1)	-0.158	-0.122
	(0.022)	(0.076)
Number of news articles	0 .004	
	(0.000)	
Return previous year * Number of news articles	-0.007	
	(0.006)	

Number of positive news articles		0 .003
		(0.151)
Number of neutral news articles		0.007
		(0.000)
Number of negative news articles		-0 .004
		(0.078)
Return previous year * Number of positive news articles		0.012
		(0.004)
Return previous year * Number of neutral news articles		-0 .021
		(0.000)
Return previous year * Number of negative news articles		-0.035
		(0.000)
Lambda	0.019	0.021
	(0.220)	(0.018)
R^2 (%)	6.93	8.57
N	4537	4537

Table D5. Relation of Fund Flows to Media Coverage considering Age of Fund

The sample includes growth funds that existed over the 1994-2000 time period. This table provides results from the Heckman two-step model for the percentage net monthly flow into the fund as the dependent variable. The independent variables are the aggregate flows to the market for all funds during the period, the lagged flow to the fund, log of the fund's total net assets under management (TNA) for the previous period, the return to the fund for the previous month, the return to the fund for the previous year, the fund's age in years, and the fund's expense ratio for the previous year.
The media coverage variables included in Model 1 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age variable and the number of news articles and between the age variable and the cumulative number of news articles. The media coverage variables included in Model 2 are the number of news articles about the fund in the period, and the cumulative number of news articles about the fund beginning in 1994 till the current period, along with the interaction terms between the age group dummy variables and the number of news articles and between the age group dummy variables and the cumulative number of news articles. p-values are provided in parentheses below the coefficient estimates.

	Model 1	**Model 2**
Intercept	0.018	-0.039
	(0.626)	(0.300)
Market Flow	0.145	0.142
	(0.000)	(0.001)
Flow (t-1)	0.003	-0.009
	(0.895)	(0.514)
Return previous month	-0.035	-0.020
	(0.000)	(0.000)
Return previous year	0.092	0.087
	(0.025)	(0.016)
Log TNA (t-1)	-0.256	-0.049
	(0.855)	(0.804)
Age	-0.001	-0.001
	(0.010)	(0.034)
Expense ratio (t-1)	-0.147	-0.150
	(0.033)	(0.029)
Number of news articles	0.005	0 .002
	(0.000)	(0.139)

Cumulative number of news articles	-0.034	-0.020
	(0.000)	(0.001)
Age * Number of news articles	-0.0001	
	(0.003)	
Age * Cumulative number of news articles	0.001	
	(0.008)	
Young funds * Number of news articles		0.012
		(0.000)
Young funds * Cumulative number of news articles		-0.011
		(0.190)
Old funds * Number of news articles		-0.001
		(0.383)
Old funds * Cumulative number of news articles		0.006
		(0.435)
Lambda	0.006	0.029
	(0.723)	(0.097)
R^2 (%)	7.35	8.52
N	4537	4505

References

Alexander, Gordon, Jonathan Jones, and Peter Nigro, 1998, Mutual fund shareholders: Characteristics, investor knowledge, and sources of information, *Financial Services Review* 7, 301-316.

Amemiya, T., 1984, Tobit models: A survey, *Journal of Econometrics* 24, 3-61.

Baks, Klaas, Andrew Metrick, and Jessica Wachter, 2001, Should investors avoid all actively managed mutual funds? A study in bayesian performance evaluation, *Journal of Finance* 56, 45-85.

Barber, Brad M., and Terrance Odean, 2004, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, University of California-Davis and University of California-Berkeley working paper.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, Out of sight, out of mind: The effects of expenses on mutual fund flows, *Journal of Business* 78, 2095-2119.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.

Berk, Jonathan and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.

Breen, Richard, 1996, Regression models: Censored, sample selected or truncated data, Thousand Oaks, Calif.: Sage Publications.

Bryant, Jennings, and Dolf Zillman, 1986, Perspectives on media effects, Lawrence Erlbaum Associates, New Jersey.

Busse, Jeff, and Clifton Green, 2002, Market efficiency in real time, *Journal of Financial Economic*s.

Chevalier, Judith and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167-1200.

Chan, Wesley, 2003, Stock price reactions to news and no-news: Drift and reversal after headlines, *Journal of Financial Economics* 70, 223-260.

Chordia, Tarun, 1996, The structure of mutual fund charges, *Journal of Financial Economics* 41, 3-39.

Cohen, Randolph, Gompers, Paul, and T. Vuolteenaho, 2002, Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and over-reactions, *Journal of Finance* 53, 1839-1885.

Das, Somnath, Carolyn B. Levine, and K. Sivaramakrishnan, 1998, Earnings predictability and bias in analysts' earnings forecasts, *The Accounting Review* 73, 277-294.

Del Guercio, Diane and Tkac, Paula, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds, *Journal of Financial and Quantitative Analysis* 37, 523-557.

Dowen, Richard J., and Thomas Mann, 2004, Mutual funds performance, management behavior, and investor costs, *Financial Services Review* 13, 79-91.

Dyck, Alexander, and Luigi Zingales, 2003, The media and asset prices, Harvard Business School and University of Chicago working paper.

Edelen, Roger, 1999, Investor flows and the assessed performance of open-end mutual funds, *Journal of Financial Economics* 53, 439-466.

Edelen, Roger, and Jerold Warner, 2001, Aggregate price effects of institutional trading: a study of mutual fund flow and market returns, *Journal of Financial Economics* 59, 195-220.

Ederington, Louis and Jae Ha Lee, 1993, How markets process information: News releases and volatility, *Journal of Finance* 48, 1161-1191.

Eveland, William P., 2001, The cognitive mediation model of learning from the news: evidence from nonelection, off-year election, and presidential election contexts, *Communication Research* 28, 571-601.

Fama, Eugene F., and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607-636.

Francis, Jennifer and Donna Philbrick, 1993, Analysts' decisions as products of multi-task environment, *Journal of Accounting Research* 31, 216-230.

Goetzmann, William, and Nadav Peles, 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research* 20, 145-158.

Greene, William H., 1981, Sample selection bias as a specification error: A comment, *Econometrica* 49, 795-798.

Greene, William H., 2000, Econometric analysis, 4th ed., Prentice Hall, New Jersey.

Griffin, Dale, and Amos Tversky, 1992, The weighing of evidence and the determinants of confidence, *Cognitive Psychology* 24, 411-435.

Gruber, Martin, 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783-810.

Harris, Richard Jackson, 2004, A cognitive psychology of mass communication, 4th ed., Mahwah, N.J. : L. Erlbaum Associates.

Heckman, James J., 1974, Shadow prices, market wages, and labor supply, *Econometrica* 42, 679-694.

Heckman, James J., 1976, The common structure of statistical models of truncation, sample selection and limited dependent variables, and a simple estimator for such models, *Annals of Economic and Social Measurement* 5, 475-492.

Heckman, James J., 1978, Dummy endogenous variables in a simultaneous equation system, *Econometrica* 46, 931-959.

Heckman, James J., 1979, Sample selection bias as a specification error, *Econometrica* 47, 153-161.

Heckman, James J., 1990, Varieties of selection bias, *American Economic Review* 80, 313-318.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *The Journal of Finance* 56, 1533-1597.

Holden, Craig, and Avanidhar Subrahmanyam, 2002, News events, information acquisition, and serial correlation, *Journal of Business* 75, 1-32.

Hong, Harrison, Terrence Lim, and Jeremy Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265-295.

Huang, Jennifer, Kelsey Wei and Hong Yan, 2004, Volatility of performance and mutual fund flows, University of Texas at Austin and Binghamton University, SUNY working paper.

Huang, Jennifer, Kelsey Wei and Hong Yan, Forthcoming, Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance*.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer: A non-event that made stock prices soar, *Journal of Finance* 56, 387–396.

Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45-70.

Jennings, Robert, and Laura Starks, 1986, Earnings announcements, stock price adjustment, and the existence of option markets, *Journal of Finance* 41, 107-125.

Kahneman, Daniel, 1973, Attention and effort, Prentice Hall, New Jersey.

Khorana, Ajay and Henri Servaes, 2003, An examination of competition in the mutual fund industry, Georgia Institute of Technology and London Business School working paper.

Khorana, Ajay and Henri Servaes, 1999, The determinants of mutual fund starts. *Review of Financial Studies* 12, 1043-1074.

Klibanoff, Peter, Owen Lamont, and Thierry Wizman, 1998, Investor reaction to salient news in closed-end country funds, *Journal of Finance* 53, 673-700.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lang, A., 2000, The limited capacity model of mediated message processing, *The Journal of Communication* 50, 46-70.

Lee, Charles M. C., 1992, Earnings news and small traders, *Journal of Accounting and Economics* 15, 265-302.

Lee, Lung-Fei, 1983, Generalized econometric models with selectivity, *Econometrica* 51, 507-512.

Lim, Terrence, 2001, Rationality and analysts' forecast bias, *Journal of Finance* 56, 369-385.

Lynch, Anthony W., and David K. Musto, 2003, How investors interpret past fund returns, *Journal of Finance* 58, 2033-2058.

Maddala, G. S., 1983, Limited-dependent and qualitative variables in econometrics, New York: Cambridge University Press.

Malhotra, D.K., and Robert W. McLeod, 1997, An empirical analysis of mutual fund expenses, *The Journal of Financial Research* 20, 175-190.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance-maximization is not the only game in town, *Journal of Financial Economics* 67, 249-304.

McQueen, Grant, and Steven Thorley, 1997, Do investors learn? Evidence from a gold market anomaly, *The Financial Review* 32, 501-525.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Mitchell, Mark L., and J. Harold Mulherin, 1994, The impact of public information on the stock market. *Journal of Finance* 49, 923-950.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, The ABCs of mutual funds: A natural experiment on fund flows and performance, University of Michigan working paper.

Newey, W., and K. West, 1987, A simple, positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Nofsinger, John, 2001, The impact of public information on investors, *Journal of Banking and Finance* 25, 1339-1366.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review*, 1279-1298.

Olson, R. J., 1980, A least squares correction for selectivity bias, *Econometrica* 48, 1815-1820.

Parse, Elizabeth M., 2001, Media effects and society, Erlbaum Associates, New Jersey.

Pastor Lubos, and Robert F. Stambaugh, 2002, Investing in equity mutual funds, *Journal of Financial Economics* 63, 351-380.

Patell, James M. and Mark A. Wolfson, 1984, The intraday speed of adjustment of stock prices to earnings and dividend announcements, *Journal of Financial Economics* 13, 223-252.

Reuter, Jonathan and Eric Zitzewitz, 2004, Do ads influence editors? Advertising and bias in the financial media, University of Oregon and Stanford Graduate School of Business working paper.

Schuster, Thomas, 2003, News events and price movements. Price effects of economic and non-economic publications in the news media, Leipzig University working paper.

Shaver, Myles J., 1998, Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival?, *Management Science* 44, 571-585.

Sirri, Erik and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589-1622.

Tetlock, Paul C., Forthcoming, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance*.

Winship, Christopher and Robert D. Mare, 1992, Models for sample selection bias, *Annual Review of Sociology* 18, 327-350.

Vita

Vasudha Vasudevan was born in Chennai, India on September 9, 1972, the daughter of M. E. Vasu and Seetha Vasu. She graduated from The Assembly of God Church School in Calcutta, India in 1989. After receiving her degree of Bachelor of Science in Microbiology from The University of Bombay in 1992, she entered the Jamnalal Bajaj Institute of Management Studies. She received the degree of Master of Management Studies from The University of Bombay in 1994. During the next four years, she was employed as an Investment banker with the Australia and New Zealand Banking Group. In 1998, she left her position as Head of Financial Market Sales for Southern India at ANZ Investment Bank and joined the doctoral program in Finance at The University of Texas at Austin.

Permanent address: 7, Ashok Arpudha, 186 Habibullah Road,
Chennai – 600017, India.

This dissertation was typed by Vasudha Vasudevan.